Jacek Banasiak

## Selected Topics in Applied Functional Analysis

October 3, 2012

## Contents

1 Basic Facts from Functional Analysis and Banach Lattices ..... 1
1.1 Spaces and Operators ..... 1
1.1.1 General Notation ..... 1
1.1.2 Operators ..... 8
1.2 Fundamental Theorems of Functional Analysis ..... 15
1.2.1 Hahn-Banach Theorem ..... 15
1.2.2 Spanning theorem and its application ..... 16
1.2.3 Banach-Steinhaus Theorem ..... 19
1.2.4 Weak compactness ..... 25
1.2.5 The Open Mapping Theorem ..... 26
1.3 Hilbert space methods ..... 28
1.3.1 To identify or not to identify-the Gelfand triple ..... 28
1.3.2 The Radon-Nikodym theorem ..... 30
1.3.3 Projection on a convex set ..... 32
1.3.4 Theorems of Stampacchia and Lax-Milgram ..... 33
1.3.5 Motivation ..... 33
1.3.6 Dirchlet problem ..... 37
1.3.7 Sobolev spaces ..... 40
1.3.8 Localization and flattening of the boundary ..... 45
1.3.9 Extension operator ..... 46
1.4 Basic applications of the density theorem ..... 50
1.4.1 Sobolev embedding ..... 50
1.4.2 Compact embedding and Rellich-Kondraschov theorem ..... 53
1.4.3 Trace theorems ..... 54
1.4.4 Regularity of variational solutions to the Dirichlet problem ..... 57
2 An Overview of Semigroup Theory ..... 65
2.1 What the semigroup theory is about ..... 65
2.2 Rudiments ..... 67
2.2.1 Definitions and Basic Properties ..... 67
2.2.2 The Hille-Yosida Theorem ..... 72
2.2.3 Dissipative operators and the Lumer-Phillips theorem ..... 75
2.2.4 Standard Examples ..... 81
2.2.5 Subspace Semigroups ..... 84
2.2.6 Sobolev Towers ..... 85
2.2.7 The Laplace Transform and the Growth Bounds of a Semigroup ..... 86
2.3 Dissipative Operators ..... 90
2.3.1 Application: Diffusion Problems ..... 92
2.3.2 Contractive Semigroups with a Parameter ..... 98
2.4 Nonhomogeneous Problems ..... 100
2.5 Positive Semigroups ..... 104
2.6 Pseudoresolvents and Approximation of Semigroups ..... 108
2.7 Uniqueness and Nonuniqueness ..... 115
References ..... 121

## Basic Facts from Functional Analysis and Banach Lattices

### 1.1 Spaces and Operators

### 1.1.1 General Notation

The symbol ':=' denotes 'equal by definition'. The sets of all natural (not including 0 ), integer, real, and complex numbers are denoted by $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$, respectively. If $\lambda \in \mathbb{C}$, then we write $\Re \lambda$ for its real part, $\Im \lambda$ for its imaginary part, and $\bar{\lambda}$ for its complex conjugate. The symbols $[a, b],(a, b)$ denote closed and open intervals in $\mathbb{R}$. Moreover,

$$
\begin{aligned}
\mathbb{R}_{+} & :=[0, \infty) \\
\mathbb{N}_{0} & :=\{0,1,2, \ldots\}
\end{aligned}
$$

If there is a need to emphasise that we deal with multidimensional quantities, we use boldface characters, for example $\mathbf{x}=\left(x_{1}, \ldots, x_{n}\right) \in \mathbb{R}^{n}$. Usually we use the Euclidean norm in $\mathbb{R}^{n}$, denoted by,

$$
|\mathbf{x}|=\sqrt{\sum_{i=1}^{n} x_{i}^{2}}
$$

If $\Omega$ is a subset of any topological space $X$, then by $\bar{\Omega}$ and Int $\Omega$ we denote, respectively, the closure and the interior of $\Omega$ with respect to $X$. If $(X, d)$ is a metric space with metric $d$, we denote by

$$
B_{x, r}:=\{y \in X ; d(x, y) \leq r\}
$$

the closed ball with centre $x$ and radius $r$. If $X$ is also a linear space, then the ball with radius $r$ centred at the origin is denoted by $B_{r}$.

Let $f$ be a function defined on a set $\Omega$ and $x \in \Omega$. We use one of the following symbols to denote this function: $f, x \rightarrow f(x)$, and $f(\cdot)$. The symbol $f(x)$ is in general reserved to denote the value of $f$ at $x$, however, occasionally we abuse this convention and use it to denote the function itself.

If $\left\{x_{n}\right\}_{n \in \mathbb{N}}$ is a family of elements of some set, then the sequence of these elements, that is, the function $n \rightarrow x_{n}$, is denoted by $\left(x_{n}\right)_{n \in \mathbb{N}}$. However, for simplicity, we often abuse this notation and use $\left(x_{n}\right)_{n \in \mathbb{N}}$ also to denote $\left\{x_{n}\right\}_{n \in \mathbb{N}}$.

The derivative operator is usually denoted by $\partial$. However, as we occasionally need to distinguish different types of derivatives of the same function, we use other commonly accepted symbols for differentiation. To indicate the variable with respect to which we differentiate we write $\partial_{t}, \partial_{x}, \partial_{t x}^{2} \ldots$. If $\mathbf{x}=\left(x_{1}, \ldots, x_{n}\right) \in \mathbb{R}^{n}$, then $\partial_{\mathbf{x}}:=\left(\partial_{x_{1}}, \ldots, \partial_{x_{n}}\right)$ is the gradient operator.

If $\beta:=\left(\beta_{1}, \ldots, \beta_{n}\right), \beta_{i} \geq 0$ is a multi-index with $|\beta|:=\beta_{1}+\cdots+\beta_{n}=k$, then symbol $\partial_{\mathbf{x}}^{\beta} f$ is any derivative of $f$ of order $k$. Thus, $\sum_{|\beta|=0}^{k} \partial^{\beta} f$ means the sum of all derivatives of $f$ of order less than or equal to $k$.

If $\Omega \subset \mathbb{R}^{n}$ is an open set, then for $k \in \mathbb{N}$ the symbol $C^{k}(\Omega)$ denotes the set of $k$ times continuously differentiable functions in $\Omega$. We denote by $C(\Omega):=C^{0}(\Omega)$ the set of all continuous functions in $\Omega$ and

$$
C^{\infty}(\Omega):=\bigcap_{k=0}^{\infty} C^{k}(\Omega)
$$

Functions from $C^{k}(\Omega)$ need not be bounded in $\Omega$. If they are required to be bounded together with their derivatives up to the order $k$, then the corresponding set is denoted by $C^{k}(\bar{\Omega})$.

For a continuous function $f$, defined on $\Omega$, we define the support of $f$ as

$$
\operatorname{supp} f=\overline{\{\mathbf{x} \in \Omega ; f(x) \neq 0\}}
$$

The set of all functions with compact support in $\Omega$ which have continuous derivatives of order smaller than or equal to $k$ is denoted by $C_{0}^{k}(\Omega)$. As above, $C_{0}(\Omega):=C_{0}^{0}(\Omega)$ is the set of all continuous functions with compact support in $\Omega$ and

$$
C_{0}^{\infty}(\Omega):=\bigcap_{k=0}^{\infty} C_{0}^{k}(\Omega)
$$

Another important standard class of spaces are the spaces $L_{p}(\Omega), 1 \leq p \leq$ $\infty$ of functions integrable with power $p$. To define them, let us establish some general notation and terminology. We begin with a measure space $(\Omega, \Sigma, \mu)$, where $\Omega$ is a set, $\Sigma$ is a $\sigma$-algebra of subsets of $\Omega$, and $\mu$ is a $\sigma$-additive measure on $\Sigma$. We say that $\mu$ is $\sigma$-finite if $\Omega$ is a countable union of sets of finite measure.

In most applications in this book, $\Omega \subset \mathbb{R}^{n}$ and $\Sigma$ is the $\sigma$-algebra of Lebesgue measurable sets. However, occasionally we need the family of Borel sets which, by definition, is the smallest $\sigma$-algebra which contains all open sets. The measure $\mu$ in the former case is called the Lebesgue measure and in the latter the Borel measure. Such measures are $\sigma$-finite.

A function $f: \Omega \rightarrow \mathbb{R}$ is said to be measurable (with respect to $\Sigma$, or with respect to $\mu$ ) if $f^{-1}(B) \in \Sigma$ for any Borel subset $B$ of $\mathbb{R}$. Because $\Sigma$ is a
$\sigma$-algebra, $f$ is measurable if (and only if) preimages of semi-infinite intervals are in $\Sigma$.

Remark 1.1. The difference between Lebesgue and Borel measurability is visible if one considers compositions of functions. Precisely, if $f$ is continuous and $g$ is measurable on $\mathbb{R}$, then $f \circ g$ is measurable but, without any additional assumptions, $g \circ f$ is not. The reason for this is that the preimage of $\{x>a\}$ through $f$ is open and preimages of open sets through Lebesgue measurable functions are measurable. On the other hand, preimage of $\{x>a\}$ through $g$ is only a Lebesgue measurable set and preimages of such sets through continuous are not necessarily measurable. To have measurability of $g \circ f$ one has to assume that preimages of sets of measure zero through $f$ are of measure zero (e.g., $f$ is Lipschitz continuous).

We identify two functions which differ from each other on a set of $\mu$ measure zero, therefore, when speaking of a function in the context of measure spaces, we usually mean a class of equivalence of functions. For most applications the distinction between a function and a class of functions is irrelevant.

One of the most important results in applications is the Łuzin theorem.
Theorem 1.2. If $f$ is Lebesgue measurable and $f(x)=0$ in the complement of a set $A$ with $\mu(A)<\infty$, then for any $\epsilon>0$ there exists a function $g \in C_{0}\left(\mathbb{R}^{n}\right)$ such that $\sup _{\mathbf{x} \in \mathbb{R}^{n}} g(\mathbf{x}) \leq \sup _{\mathbf{x} \in \mathbb{R}^{n}} f(\mathbf{x})$ and $\mu(\{\mathbf{x} ; f(\mathbf{x}) \neq g(\mathbf{x})\})<\epsilon$.

In other words, the theorem implies that there is a sequence of compactly supported continuous functions converging to $f$ almost everywhere. Indeed, for any $n$ we find a continuous function $\phi_{n}$ such that for $A_{n}=\left\{\mathbf{x} ; \phi_{n}(\mathbf{x}) \neq f(\mathbf{x})\right\}$ we have

$$
\mu\left(A_{n}\right) \leq \frac{1}{n^{2}}
$$

Define

$$
A=\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_{n} .
$$

We see that if $\mathbf{x} \notin A$, then there is $k$ such that for any $n \geq k, \mathbf{x} \notin A_{k}$, that is, $\phi_{n}(\mathbf{x})=f(\mathbf{x})$ and hence $\phi_{n}(\mathbf{x}) \rightarrow f(\mathbf{x})$ whenever $\mathbf{x} \notin A$. On the other hand,

$$
0 \leq \mu(A) \leq \lim _{k \rightarrow \infty} \sum_{n=k}^{\infty} \frac{1}{n^{2}}=0
$$

and hence $\left(\phi_{n}\right)_{n \in \mathbb{N}}$ converges to $f$ almost everywhere.
The space of equivalence classes of all measurable real functions on $\Omega$ is denoted by $L_{0}(\Omega, d \mu)$ or simply $L_{0}(\Omega)$.

The integral of a measurable function $f$ with respect to measure $\mu$ over a set $\Omega$ is written as

$$
\int_{\Omega} f d \mu=\int_{\Omega} f(\mathbf{x}) d \mu_{\mathbf{x}}
$$

where the second version is used if there is a need to indicate the variable of integration. If $\mu$ is the Lebesgue measure, we abbreviate $d \mu_{\mathbf{x}}=d \mathbf{x}$.

For $1 \leq p<\infty$ the spaces $L_{p}(\Omega)$ are defined as subspaces of $L_{0}(\Omega)$ consisting of functions for which

$$
\begin{equation*}
\|f\|_{p}:=\|f\|_{L_{p}(\Omega)}=\left(\int_{\Omega}|f(\mathbf{x})|^{p} d \mathbf{x}\right)^{1 / p}<\infty \tag{1.1}
\end{equation*}
$$

The space $L_{p}(\Omega)$ with the above norm is a Banach space. It is customary to complete the scale of $L_{p}$ spaces by the space $L_{\infty}(\Omega)$ defined to be the space of all Lebesgue measurable functions which are bounded almost everywhere in $\Omega$, that is, bounded everywhere except possibly on a set of measure zero. The corresponding norm is defined by

$$
\begin{equation*}
\|f\|_{\infty}:=\|f\|_{L_{\infty}(\Omega)}:=\inf \{M ; \mu(\{\mathbf{x} \in \Omega ;|f(\mathbf{x})|>M\})=0\} \tag{1.2}
\end{equation*}
$$

The expression on the right-hand side of (1.2) is frequently referred to as the essential supremum of $f$ over $\Omega$ and denoted $\operatorname{ess}_{\sup }^{\mathbf{x} \in \Omega} \boldsymbol{}|f(\mathbf{x})|$.

If $\mu(\Omega)<\infty$, then for $1 \leq p \leq p^{\prime} \leq \infty$ we have

$$
\begin{equation*}
L_{p^{\prime}}(\Omega) \subset L_{p}(\Omega) \tag{1.3}
\end{equation*}
$$

and for $f \in L_{\infty}(\Omega)$

$$
\begin{equation*}
\|f\|_{\infty}=\lim _{p \rightarrow \infty}\|f\|_{p} \tag{1.4}
\end{equation*}
$$

which justifies the notation. However,

$$
\bigcap_{1 \leq p<\infty} L_{p}(\Omega) \neq L_{\infty}(\Omega)
$$

as demonstrated by the function $f(x)=\ln x, x \in(0,1]$. If $\mu(\Omega)=\infty$, then neither (1.3) nor (1.4) hold.

Occasionally we need functions from $L_{0}(\Omega)$ which are $L_{p}$ only on compact subsets of $\mathbb{R}^{n}$. Spaces of such functions are denoted by $L_{p, l o c}(\Omega)$. A function $f \in L_{1, l o c}(\Omega)$ is called locally integrable (in $\Omega$ ).

Let $\Omega \subset \mathbb{R}^{n}$ be an open set. It is clear that

$$
C_{0}^{\infty}(\Omega) \subset L_{p}(\Omega)
$$

for $1 \leq p \leq \infty$. If $p \in[1, \infty)$, then we have even more: $C_{0}^{\infty}(\Omega)$ is dense in $L_{p}(\Omega)$.

$$
\begin{equation*}
\overline{C_{0}^{\infty}(\Omega)}=L_{p}(\Omega) \tag{1.5}
\end{equation*}
$$

where the closure is taken in the $L_{p}$-norm.

Example 1.3. Having in mind further applications, it is worthwhile to have some understanding of the structure of this result; see [4, Lemma 2.18]. Let us define the function

$$
\omega(\mathbf{x})= \begin{cases}\exp \left(\frac{1}{|\mathbf{x}|^{2}-1}\right) & \text { for }|\mathbf{x}|<1  \tag{1.6}\\ 0 & \text { for }|\mathbf{x}| \geq 1\end{cases}
$$

This is a $C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ function with support $B_{1}$.
Using this function we construct the family

$$
\omega_{\epsilon}(\mathbf{x})=C_{\epsilon} \omega(\mathbf{x} / \epsilon)
$$

where $C_{\epsilon}$ are constants chosen so that $\int_{\mathbb{R}^{n}} \omega_{\epsilon}(\mathbf{x}) d \mathbf{x}=1$; these are also $C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ functions with support $B_{\epsilon}$, often referred to as mollifiers. Using them, we define the regularisation (or mollification) of $f$ by taking the convolution

$$
\begin{equation*}
\left(J_{\epsilon} * f\right)(\mathbf{x}):=\int_{\mathbb{R}^{n}} f(\mathbf{x}-\mathbf{y}) \omega_{\epsilon}(\mathbf{y}) d \mathbf{y}=\int_{\mathbb{R}^{n}} f(\mathbf{y}) \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{y} \tag{1.7}
\end{equation*}
$$

Precisely speaking, if $\Omega \neq \mathbb{R}^{n}$, we integrate outside the domain of definition of $f$. Thus, in such cases below, we consider $f$ to be extended by 0 outside $\Omega$.

Then, we have
Theorem 1.4. With the notation above,

1. Let $p \in[1, \infty)$. If $f \in L_{p}(\Omega)$, then

$$
\lim _{\epsilon \rightarrow 0^{+}}\left\|J_{\epsilon} * f-f\right\|_{p}=0
$$

2. If $f \in C(\Omega)$, then $J_{\epsilon} * f \rightarrow f$ uniformly on any $\bar{G} \subset \Omega$.
3. If $\bar{\Omega}$ is compact and $f \in C(\bar{\Omega})$, then $J_{\epsilon} * f \rightarrow f$ uniformly on $\bar{\Omega}$.

Proof. For 1.-3., even if $\mu(\Omega)=\infty$, then any $f \in L_{p}(\Omega)$ can be approximated by (essentially) bounded (simple) functions with compact supports. It is enough to consider a real nonnegative function $u$. For such a $u$, there is a monotonically increasing sequence $\left(s_{n}\right)_{n \in \mathbb{N}}$ of nonnegative simple functions converging point-wise to $u$ on $\Omega$. Since $0 \leq s_{n}(\mathbf{x}) \leq u(\mathbf{x})$, we have $s_{n} \in L_{p}(\Omega)$, $\left(u(\mathbf{x})-s_{n}(\mathbf{x})\right)^{p} \leq u^{p}(\mathbf{x})$ and thus $s_{n} \rightarrow u$ in $L_{p}(\Omega)$ by the Dominated Convergence Theorem. Thus there exists a function $s$ in the sequence for which $\|u-s\|_{p} \leq \epsilon / 2$. Since $p<\infty$ and $s$ is simple, the support of $s$ must have finite volume. We can also assume that $s(\mathbf{x})=0$ outside $\Omega$. By the Luzin theorem, there is $\phi \in C_{0}\left(\mathbb{R}^{n}\right)$ such that $|\phi(\mathbf{x})| \leq\|s\|_{\infty}$ for all $\mathbf{x} \in \mathbb{R}^{n}$ and

$$
\mu\left(\left\{\mathbf{x} \in \mathbb{R}^{n} ; \phi(\mathbf{x}) \neq s(\mathbf{x})\right\} \leq\left(\frac{\epsilon}{4\|s\|_{\infty}}\right)^{p}\right.
$$

Hence

$$
\|s-\phi\|_{p} \leq\|s-\phi\|_{\infty} \frac{\epsilon}{4\|s\|_{\infty}} \leq \frac{\epsilon}{2}
$$

and $\|u-\phi\|_{p}<\epsilon$.
Therefore, first we prove the result for continuous compactly supported functions.

Because the effective domain of integration in the second integral is $B_{\mathbf{x}, \epsilon}$, $J_{\epsilon} * f$ is well defined whenever $f$ is locally integrable and, similarly, if the support of $f$ is bounded, then $\operatorname{supp} J_{\epsilon} * f$ is also bounded and it is contained in the $\epsilon$-neighbourhood of $\operatorname{supp} f$. The functions $f_{\epsilon}$ are infinitely differentiable with

$$
\begin{equation*}
\partial_{\mathbf{x}}^{\beta}\left(J_{\epsilon} * f\right)(\mathbf{x})=\int_{\mathbb{R}^{n}} f(\mathbf{y}) \partial_{\mathbf{x}}^{\beta} \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{y} \tag{1.8}
\end{equation*}
$$

for any $\beta$. By Hölder inequality, if $f \in L_{p}\left(\mathbb{R}^{n}\right)$, then $J_{\epsilon} * f \in L_{p}\left(\mathbb{R}^{n}\right)$ with

$$
\begin{equation*}
\left\|J_{\epsilon} * f\right\|_{p} \leq\|f\|_{p} \tag{1.9}
\end{equation*}
$$

for any $\epsilon>0$. Indeed, for $p=1$

$$
\int_{\mathbb{R}^{n}}\left|J_{\epsilon} * f(\mathbf{x})\right| d \mathbf{x} \leq \int_{\mathbb{R}^{n}}|f(\mathbf{y})|\left(\int_{\mathbb{R}^{n}} \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{x}\right) d \mathbf{y}=\|f\|_{1}
$$

For $p>1$, we have

$$
\begin{aligned}
\left|J_{\epsilon} * f(\mathbf{x})\right| & =\left|\int_{\mathbb{R}^{n}} f(\mathbf{y}) \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{y}\right| \\
& \leq\left(\int_{\mathbb{R}^{n}} \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{y}\right)^{1 / q}\left(\int_{\mathbb{R}^{n}}|f(\mathbf{y})|^{p} \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{y}\right)^{1 / p} \\
& =\left(\int_{\mathbb{R}^{n}}|f(\mathbf{y})|^{p} \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{y}\right)^{1 / p}
\end{aligned}
$$

and, as above,

$$
\int_{\mathbb{R}^{n}}\left|J_{\epsilon} * f(\mathbf{x})\right|^{p} d \mathbf{x} \leq \int_{\mathbb{R}^{n}}|f(\mathbf{y})|^{p}\left(\int_{\mathbb{R}^{n}} \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{x}\right) d \mathbf{y}=\|f\|_{p}^{p}
$$

Next (remember $f$ is compactly supported continuous function, and thus it is uniformly continuous)

$$
\begin{aligned}
\left|\left(J_{\epsilon} * f\right)(\mathbf{x})-f(\mathbf{x})\right| & =\left|\int_{\mathbb{R}^{n}} f(\mathbf{y}) \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{y}-\int_{\mathbb{R}^{n}} f(\mathbf{x}) \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{y}\right| \\
& \leq \int_{\mathbb{R}^{n}}|f(\mathbf{y})-f(\mathbf{x})| \omega_{\epsilon}(\mathbf{x}-\mathbf{y}) d \mathbf{y} \leq \sup _{\|\mathbf{x}-\mathbf{y}\| \leq \epsilon}|f(\mathbf{x})-f(\mathbf{y})| .
\end{aligned}
$$

By the compactness of support, and thus uniform continuity, of $f$ we obtain $J_{\epsilon} * f \rightrightarrows f$ and, again by compactness of the support,

$$
\begin{equation*}
f=\lim _{\epsilon \rightarrow 0^{+}} f_{\epsilon} \quad \text { in } L_{p}\left(\mathbb{R}^{n}\right) \tag{1.10}
\end{equation*}
$$

as well as in $C(\bar{\Omega})$, where in the latter case we extend $f$ outside $\Omega$ by a continuous function (e.g. by the Urysohn theorem).

To extend the result to an arbitrary $f \in L_{p}(\Omega)$, let $\phi \in C_{0}(\Omega)$ such that $\|f-\phi\|_{p}<\eta$ and $\left\|J_{\epsilon} * \phi-\phi\right\|_{p}<\eta$

$$
\begin{aligned}
\left\|J_{\epsilon} * f-f\right\|_{p} & \left.\leq \| J_{\epsilon} * f-J_{\epsilon} * \phi\right)\left\|_{p}+\right\| J_{\epsilon} * \phi-\phi\left\|_{p}+\right\| f-\phi \| \\
& \leq 2\|f-\phi\|+\left\|J_{\epsilon} * \phi-\phi\right\|_{p}<\eta
\end{aligned}
$$

for sufficiently small $\epsilon$.
As an example of application, we shall consider a generalization of the duBois-Reymond lemma. Let $\Omega \subset \mathbb{R}^{n}$ be an open set and let $u \in L_{1, l o c}(\Omega)$ be such that

$$
\int_{\Omega} u(\mathbf{x}) f(\mathbf{x}) d \mathbf{x}=0
$$

for any $C_{0}^{\infty}(\Omega)$. Then $u=0$ almost everywhere on $\Omega$. To prove this statement, let $g \in L_{\infty}(\Omega)$ such that $\operatorname{supp} g$ is a compact set in $\Omega$. We define $g_{m}=J_{1 / m} * g$. Then $g_{m} \in C_{0}^{\infty}(\Omega)$ for large $m$. Since a compactly supported bounded function is integrable, we have $g_{m} \rightarrow g$ in $L_{1}(\Omega)$ and thus there is a subsequence (denoted by the same indices) such that $g_{m} \rightarrow g$ almost everywhere. Moreover, $\left\|g_{m}\right\|_{\infty} \leq\|g\|_{\infty}$. Using compactness of the supports and dominated convergence theorem, we obtain

$$
\int_{\Omega} u(\mathbf{x}) g(\mathbf{x}) d \mathbf{x}=0
$$

If we take any compact set $K \subset \Omega$ and define $g=\operatorname{sign} u$ on $K$ and 0 otherwise, we find that for any $K$,

$$
\int_{K}|u(\mathbf{x})| d \mathbf{x}=0
$$

Hence $u=0$ almost everywhere on $K$ and, since $K$ was arbitrary, this holds almost everywhere on $\Omega$.

Remark 1.5. We observe that, if $f$ is nonnegative, then $f_{\epsilon}$ are also nonnegative by (1.7) and hence any non-negative $f \in L_{p}\left(\mathbb{R}^{n}\right)$ can be approximated by nonnegative, infinitely differentiable, functions with compact support.

Remark 1.6. Spaces $L_{p}(\Omega)$ often are defined as a completion of $C_{0}(\Omega)$ in the $L_{p}(\Omega)$ norm, thus avoiding introduction of measure theory. The theorem above shows that these two definitions are equivalent.

### 1.1.2 Operators

Let $X, Y$ be real or complex Banach spaces with the norm denoted by $\|\cdot\|$ or $\|\cdot\|_{X}$.

An operator from $X$ to $Y$ is a linear rule $A: D(A) \rightarrow Y$, where $D(A)$ is a linear subspace of $X$, called the domain of $A$. The set of operators from $X$ to $Y$ is denoted by $L(X, Y)$. Operators taking their values in the space of scalars are called functionals. We use the notation $(A, D(A))$ to denote the operator $A$ with domain $D(A)$. If $A \in L(X, X)$, then we say that $A$ (or $(A, D(A))$ ) is an operator in $X$.

By $\mathcal{L}(X, Y)$, we denote the space of all bounded operators between $X$ and $Y ; \mathcal{L}(X, X)$ is abbreviated as $\mathcal{L}(X)$. The space $\mathcal{L}(X, Y)$ can be made a Banach space by introducing the norm of an operator $X$ by

$$
\begin{equation*}
\|A\|=\sup _{\|x\| \leq 1}\|A x\|=\sup _{\|x\|=1}\|A x\| . \tag{1.11}
\end{equation*}
$$

If $(A, D(A))$ is an operator in $X$ and $Y \subset X$, then the part of the operator $A$ in $Y$ is defined as

$$
\begin{equation*}
A_{Y} y=A y \tag{1.12}
\end{equation*}
$$

on the domain

$$
D\left(A_{Y}\right)=\{x \in D(A) \cap Y ; \quad A x \in Y\}
$$

A restriction of $(A, D(A))$ to $D \subset D(A)$ is denoted by $\left.A\right|_{D}$. For $A, B \in$ $L(X, Y)$, we write $A \subset B$ if $D(A) \subset D(B)$ and $\left.B\right|_{D(A)}=A$.

Two operators $A, B \in \mathcal{L}(X)$ are said to commute if $A B=B A$. It is not easy to extend this definition to unbounded operators due to the difficulties with defining the domains of the composition. The extension is usually done to the case when one of the operators is bounded. Thus, an operator $A \in L(X)$ is said to commute with $B \in \mathcal{L}(X)$ if

$$
\begin{equation*}
B A \subset A B \tag{1.13}
\end{equation*}
$$

This means that for any $x \in D(A), B x \in D(A)$ and $B A x=A B x$.
We define the image of $A$ by

$$
\operatorname{Im} A=\{y \in Y ; y=A x \text { for some } x \in D(A)\}
$$

and the kernel of $A$ by

$$
\operatorname{Ker} A=\{x \in D(A) ; A x=0\} .
$$

We note a simple result which is frequently used throughout the book.
Proposition 1.7. Suppose that $A, B \in L(X, Y)$ satisfy: $A \subset B, \operatorname{Ker} B=\{0\}$, and $\operatorname{Im} A=Y$. Then $A=B$.

Proof. If $D(A) \neq D(B)$, we take $x \in D(B) \backslash D(A)$ and let $y=B x$. Because $A$ is onto, there is $x^{\prime} \in D(A)$ such that $y=A x^{\prime}$. Because $x^{\prime} \in D(A) \subset D(B)$ and $A \subset B$, we have $y=A x^{\prime}=B x^{\prime}$ and $B x^{\prime}=B x$. Because $\operatorname{Ker} B=\{0\}$, we obtain $x=x^{\prime}$ which is a contradiction with $x \notin D(A)$.

Furthermore, the graph of $A$ is defined as

$$
\begin{equation*}
G(A)=\{(x, y) \in X \times Y ; x \in D(A), y=A x\} \tag{1.14}
\end{equation*}
$$

We say that the operator $A$ is closed if $G(A)$ is a closed subspace of $X \times Y$. Equivalently, $A$ is closed if and only if for any sequence $\left(x_{n}\right)_{n \in \mathbb{N}} \subset D(A)$, if $\lim _{n \rightarrow \infty} x_{n}=x$ in $X$ and $\lim _{n \rightarrow \infty} A x_{n}=y$ in $Y$, then $x \in D(A)$ and $y=A x$.

An operator $A$ in $X$ is closable if the closure of its graph $\overline{G(A)}$ is itself a graph of an operator, that is, if $(0, y) \in \overline{G(A)}$ implies $y=0$. Equivalently, $A$ is closable if and only if for any sequence $\left(x_{n}\right)_{n \in \mathbb{N}} \subset D(A)$, if $\lim _{n \rightarrow \infty} x_{n}=0$ in $X$ and $\lim _{n \rightarrow \infty} A x_{n}=y$ in $Y$, then $y=0$. In such a case the operator whose graph is $\overline{G(A)}$ is called the closure of $A$ and denoted by $\bar{A}$.

By definition, when $A$ is closable, then

$$
\begin{aligned}
& D(\bar{A})=\left\{x \in X ; \text { there is }\left(x_{n}\right)_{n \in \mathbb{N}} \subset D(A) \text { and } y \in X\right. \text { such that } \\
& \left.\qquad \quad\left\|x_{n}-x\right\| \rightarrow 0 \text { and }\left\|A x_{n}-y\right\| \rightarrow 0\right\} \\
& \bar{A} x=y .
\end{aligned}
$$

For any operator $A$, its domain $D(A)$ is a normed space under the graph norm

$$
\begin{equation*}
\|x\|_{D(A)}:=\|x\|_{X}+\|A x\|_{Y} \tag{1.15}
\end{equation*}
$$

The operator $A: D(A) \rightarrow Y$ is always bounded with respect to the graph norm, and $A$ is closed if and only if $D(A)$ is a Banach space under (1.15).

## The differentiation operator

One of the simplest and most often used unbounded, but closed or closable, operators is the operator of differentiation. If $X$ is any of the spaces $C([0,1])$ or $L_{p}([0,1])$, then considering $f_{n}(x):=C_{n} x^{n}$, where $C_{n}=1$ in the former case and $C_{n}=(n p+1)^{1 / p}$ in the latter, we see that in all cases $\left\|f_{n}\right\|=1$. However,

$$
\begin{gathered}
\|\overbrace{n}\|_{p}\left(Q_{1}\right) \\
\operatorname{Lin}_{n}^{\prime} \|=n\left(\frac{n p+1}{n p+1-p}\right)^{1 / p}
\end{gathered}
$$

$$
T=\overbrace{}^{1}
$$


in $L_{p}([0,1])$ and $\left\|f_{n}^{\prime}\right\|=n$ in $C([0,1])$, so that the operator of differentiation is unbounded.

Let us define $T f=f^{\prime}$ as an unbounded operator on $D(T)=\{f \in X ; T f \in$ $X\}$, where $X$ is any of the above spaces. We can easily see that in $X=C([0,1])$ the operator $T$ is closed. Indeed, let us take $\left(f_{n}\right)_{n \in \mathbb{N}}$ such that $\lim _{n \rightarrow \infty} f_{n}=f$ and $\lim _{n \rightarrow \infty} T f_{n}=g$ in $X$. This means that $\left(f_{n}\right)_{n \in \mathbb{N}}$ and $\left(f_{n}^{\prime}\right)_{n \in \mathbb{N}}$ converge uniformly to, respectively, $f$ and $g$, and from basic calculus $f$ is differentiable and $f^{\prime}=g$.

The picture changes, however, in $L_{p}$ spaces. To simplify the notation, we take $p=1$ and consider the sequence of functions

$$
f_{n}(x)= \begin{cases}0 & \text { for } 0 \leq x \leq \frac{1}{2}, \\ \frac{n}{2}\left(x-\frac{1}{2}\right)^{2} & \text { for } \frac{1}{2}<x \leq \frac{1}{2}+\frac{1}{n} \\ x-\frac{1}{2}-\frac{1}{2 n} & \text { for } \frac{1}{2}+\frac{1}{n}<x \leq 1\end{cases}
$$

These are differentiable functions and it is easy to see that $\left(f_{n}\right)_{n \in \mathbb{N}}$ converges in $L_{1}([0,1])$ to the function $f$ given by $f(x)=0$ for $x \in[0,1 / 2]$ and $f(x)=$ $x-1 / 2$ for $x \in(1 / 2,1]$ and the derivatives converge to $g(x)=0$ if $x \in[0,1 / 2]$
$\frac{1}{2} \quad \frac{1}{2} \times \frac{1}{n}$
 and to $g(x)=1$ otherwise. The function $f$, however, is not differentiable and so $T$ is not closed. On the other hand, $g$ seems to be a good candidate for the derivative of $f$ in some more general sense. Let us develop this idea further. First, we show that $T$ is closable. Let $\left(f_{n}\right)_{n \in \mathbb{N}}$ and $\left(f_{n}^{\prime}\right)_{n \in \mathbb{N}}$ converge in $X$ to $f$ and $g$, respectively. Then, for any $\phi \in C_{0}^{\infty}((0,1))$, we have, integrating by parts,

$$
\int_{0}^{1} f_{n}^{\prime}(x) \phi(x) d x=-\int_{0}^{1} f_{n}(x) \phi^{\prime}(x) d x
$$

and because we can pass to the limit on both sides, we obtain

$$
\begin{equation*}
\int_{0}^{1} g(x) \phi(x) d x=-\int_{0}^{1} f(x) \phi^{\prime}(x) d x \tag{1.16}
\end{equation*}
$$

Using the equivalent characterization of closability, we put $f=0$, so that

$$
\int_{0}^{1} g(x) \phi(x) d x=0
$$


for any $\phi \in C_{0}^{\infty}((0,1))$ which yields $g(x)=0$ almost everywhere on $[0,1]$. Hence $g=0$ in $L_{1}([0,1])$ and consequently $T$ is closable.

The domain of $\bar{T}$ in $L_{1}([0,1])$ is called the Sobolev space $W_{1}^{1}([0,1])$ which is discussed in more detail in Subsection 2.3.1.

These considerations can be extended to hold in any $\Omega \subset \mathbb{R}^{n}$. In particular, we can use (1.16) to generalize the operation of differentiation in the following


1.1 Spaces and Operators
way: we say that a function $g \in L_{1, l o c}(\Omega)$ is the generalised (or distributional) derivative of $f \in L_{1, l o c}(\Omega)$ of order $\alpha$, denoted by $\partial_{\mathbf{x}}^{\alpha} f$, if

$$
\begin{equation*}
\int_{\Omega} g(\mathbf{x}) \phi(\mathbf{x}) d \mathbf{x}=(-1)^{|\beta|} \int_{\Omega} f(\mathbf{x}) \partial_{\mathbf{x}}^{\beta} \phi(\mathbf{x}) d \mathbf{x} \tag{1.17}
\end{equation*}
$$

for any $\phi \in C_{0}^{\infty}(\Omega)$.
This operation is well defined. This follows from the du Bois Reymond lemma.

From the considerations above it is clear that $\partial_{\mathbf{x}}^{\beta}$ is a closed operator extending the classical differentiation operator (from $C^{|\beta|}(\Omega)$ ). One can also prove that $\partial_{\mathbf{x}}^{\beta}$ is the closure of the classical differentiation operator.
Proposition 1.8. If $\Omega=\mathbb{R}^{n}$, then $\partial_{\mathbf{x}}^{\beta}$ is the closure of the classical differentiation operator.

Proof. We use (1.7) and (1.8). Indeed, let $f \in L_{p}\left(\mathbb{R}^{n}\right)$ and $g=D^{\alpha} f \in L_{p}\left(\mathbb{R}^{n}\right)$. We consider $f_{\epsilon}=J_{\epsilon} * f \rightarrow f$ in $L_{p}$. By the Fubini theorem, we prove

$$
\begin{aligned}
\int_{\mathbb{R}^{n}}\left(J_{\epsilon} * f\right)(\mathbf{x}) D^{\alpha} \phi(\mathbf{x}) d \mathbf{x} & =\int_{\mathbb{R}^{n}} \omega_{\epsilon}(y) \int_{\mathbb{R}^{n}} f(\mathbf{x}-\mathbf{y}) D^{\alpha} \phi(\mathbf{x}) d \mathbf{x} d \mathbf{y} \\
& =(-1)^{|\alpha|} \int_{\mathbb{R}^{n}} \omega_{\epsilon}(y) \int_{\mathbb{R}^{n}} g(\mathbf{x}-\mathbf{y}) \phi(\mathbf{x}) d \mathbf{x} d \mathbf{y} \\
& =(-1)^{|\alpha|} \int_{\mathbb{R}^{n}}\left(J_{\epsilon} * g\right) \phi(\mathbf{x}) d \mathbf{x}
\end{aligned}
$$

so that $D^{\alpha} f_{\epsilon}=J_{\epsilon} * D^{\alpha} f=J_{\epsilon} * g \rightarrow g$ as $\epsilon \rightarrow 0$ in $L_{p}$. This shows that action of the distributional derivative can be obtained as the closure of the classical derivation.

Otherwise the proof is more complicated (see, e.g., [4, Theorem 3.16]) since we do not know whether we can extend $f$ outside $\Omega$ in such a way that the extension still will have the generalized derivative. We shall discuss it later.

Example 1.9. A non closable operator. Let us consider the space $X=$ $L_{\mathbb{1}}((0,1))$ and the operator $K: X \rightarrow Y, Y=X \times \mathbb{C}$ (with the Euclidean norm), defined by

$$
\begin{equation*}
K v=<v, v(1)> \tag{1.18}
\end{equation*}
$$

on the domain $D(K)$ consisting of continuous functions on $[0,1]$. We have the following lemma

Lemma 1.10. $K$ is not closable, but has a bounded inverse. ImK is dense in $Y$.

Proof. Let $f \in C^{\infty}([0,1])$ be such that

$$
f(x)=\left\{\begin{array}{l}
0 \text { for } 0 \leq x<1 / 3 \\
1 \text { for } 2 / 3<x \leq 1
\end{array}\right.
$$

To construct such a function, we can consider e.g. $J_{1 / 3} * \bar{f}$ where

$$
\bar{f}(x)=\left\{\begin{array}{l}
1 \quad \text { for } \quad 2 / 3<x \leq 1 \\
0 \text { otherwise }
\end{array}\right.
$$

Let $v_{n}(x)=f\left(x^{n}\right)$ for $0 \leq x \leq 1$. Clearly, $v_{n} \in D(K)$ and $v_{n} \rightarrow 0$ in $L_{2}((0,1))$ as

$$
\int_{0}^{1} f^{2}\left(x^{n}\right) d x=\int_{3-1 / n}^{1} f^{2}\left(x^{n}\right) d x=\frac{1}{n} \int_{1 / 3}^{1} z^{-1+1 / n} f^{2}(z) d z
$$

However, $K v_{n}=<v_{n}, 1>\rightarrow<0,1>\neq<0,0>$.
Further, $K$ is one-to-one with $K^{-1}(v, v(1))=v$ and

$$
\left\|K^{-1}(v, v(1))\right\|^{2}=\|v\|^{2} \leq\|v\|^{2}+|v(1)|^{2}
$$

To prove that $\operatorname{Im} K$ is dense in $Y$, let $<y, \alpha>\in Y$. We know that $C_{0}^{\infty}((0,1)) \subset D(K)$ is dense in $Z=L_{2}((0,1))$. Let $\left(\phi_{n}\right)$ be sequence of $C_{0}^{\infty}$ functions which approximate $y$ in $L_{2}(0,1)$ and put $w_{n}=\phi_{n}+\alpha v_{n}$. We have $K w_{n}=<w_{n}, \alpha>\rightarrow<y, \alpha>$.

## Absolutely continuous functions

In one-dimensional spaces the concept of the generalised derivative is closely related to a classical notion of absolutely continuous function. Let $I=[a, b] \subset$ $\mathbb{R}^{1}$ be a bounded interval. We say that $f: I \rightarrow \mathbb{C}$ is absolutely continuous if, for any $\epsilon>0$, there is $\delta>0$ such that for any finite collection $\left\{\left(a_{i}, b_{i}\right)\right\}_{i}$ of disjoint intervals in $[a, b]$ satisfying $\sum_{i}\left(b_{i}-a_{i}\right)<\delta$, we have $\sum_{i}\left|f\left(b_{i}\right)-f\left(a_{i}\right)\right|<\epsilon$. The fundamental theorem of calculus, [150, Theorem 8.18], states that any absolutely continuous function $f$ is differentiable almost everywhere, its derivative $f^{\prime}$ is Lebesgue integrable on $[a, b]$, and $f(t)-f(a)=\int_{a}^{t} f^{\prime}(s) d s$. It can be proved (e.g., [61, Theorem VIII.2]) that absolutely continuous functions on $[a, b]$ are exactly integrable functions having integrable generalised derivatives and the generalised derivative of $f$ coincides with the classical derivative of $f$ almost everywhere.

Let us explore this connection. We prove
Theorem 1.11. Assume that $u \in L_{1, \text { loc }}(\mathbb{R})$ and its generalized derivative $D u$ also satisfies $D u \in L_{1, \text { loc }}(\mathbb{R})$. Then there is a continuous representation $\widetilde{u}$ of u such that

$$
\widetilde{u}(x)=C+\int_{0}^{x} D u(t) d t
$$

for some constant $C$ and thus $u$ is differentiable almost everywhere.

Proof. The proof is carried out in three steps. In Step 1, we prove that if

$$
\begin{equation*}
F(x)=\int_{a}^{x} f(y) d y \tag{1.19}
\end{equation*}
$$

where $f \in L_{1, \text { loc }}(\mathbb{R})$, then $F$ is differentiable almost everywhere (it is absolutely continuous) and $f$ is its derivative. In Step 2, we show that if an $L_{1, l o c}(\mathbb{R})$ function has the generalised derivative equal to zero, then it is constant (almost everywhere). Finally, in Step 3, we show that the generalised derivative of $F$ defined by (1.19) coincides with $f$, which will allow to draw the final conclusion.

Step 1. Consider

$$
A_{h} f(x)=\frac{1}{h} \int_{x}^{x+h} f(y) d y
$$

Clearly it is a jointly continuous function on $\mathbb{R}_{+} \times \mathbb{R}$. Further, denote

$$
H f(x)=\sup _{h>0}\left|A_{h} f(x)\right|
$$

We restrict considerations to some bounded open interval $I$. Then $A_{h} f(x) \rightarrow$ $f(x)$ if there is no $n$ such that $x \in S_{n}=\left\{x ; \limsup _{h \rightarrow 0}\left|A_{h} f(x)-f(x)\right| \geq\right.$ $1 / n\}$. Thus, we have to prove $\mu\left(S_{n}\right)=0$ for any $n$.

Then we can assume that $f$ is of bounded support and therefore, by the Luzin theorem, for any $\epsilon$ there is a continuous function $g$ with bounded support with $\mu(\{x \in I, f(x) \neq g(x)\}) \leq \epsilon$. Fix any $\epsilon$ and corresponding $g$. Then

$$
\limsup _{h \rightarrow 0}\left|A_{h} f(x)-f(x)\right| \leq \sup _{h>0}\left|A_{h}(f(x)-g(x))\right|+\lim _{h \rightarrow 0}\left|A_{h} g(x)-g(x)\right|+|f(x)-g(x)|
$$

The second term is zero, the third is 0 outside a set of measure $\epsilon$. We need to estimate the first term. For a given $\phi$ consider an open set $E_{\alpha}=\{x \in$ $I ; H \phi(x)>\alpha\}$. For any $x \in E_{\alpha}$ we find $I_{x, r_{x}}=\left(x-r_{x}, x+r_{x}\right)$ such that $A_{h} f(x)>\alpha$. Thus, $E_{\alpha}$ is covered by these intervals. From the theory of Lebesque measure, the measure of any measurable set $S$ is supremum over measures of compact sets $K \subset S$. Thus, for any $c<\mu\left(E_{\alpha}\right)$ we can find compact set $K \subset E_{\alpha}$ with $c<\mu(K) \subset \mu\left(E_{\alpha}\right)$ and a finite cover of $K$ by $I_{x_{i}, r_{x_{i}}}, i=1, \ldots, i_{K}$. Let us modify this cover in the following way. Let $I_{1}$ be the element of maximum length $2 r_{1}, I_{2}$ be the largest of the remaining which are disjoint with $I_{1}$ and so on, until the collection is exhausted with $j=J$. According to the construction, if some $I_{x_{i}, r_{x_{i}}}$ is not in the selected list, then there is $j$ such that $I_{x_{i}, r_{x_{i}}} \cap I_{j} \neq \emptyset$. Let as take the smallest such $j$, that is, the largest $I_{j}$. Then $2 r_{x_{i}}$ is at most equal to the length of $I_{j}, 2 r_{j}$, and thus $I_{x_{i}, r_{x_{i}}} \subset I_{j}^{*}$ where the latter is the interval with the same centre as $I_{j}$ but with length $6 r_{j}$. The collection of $I_{j}^{*}$ also covers $K$ and we have

$$
c \leq 6 \sum_{j=1}^{J} r_{j}=3 \sum_{j=1}^{J} \mu\left(I_{j}\right) \leq \frac{3}{\alpha} \sum_{j=1}^{J} \int_{I_{j}}|\phi(y)| d y \leq \frac{3}{\alpha} \int_{I}|\phi(y)| d y
$$

Passing with $c \rightarrow \mu\left(E_{\alpha}\right)$ we get

$$
\mu\left(E_{\alpha}\right)=\mu(\{x \in I ; H \phi(x)>\alpha\}) \leq \frac{3}{\alpha} \int_{I}|\phi(y)| d y .
$$

Using this for $\phi=f-g$ we see that for any $\epsilon>0$ we have

$$
\mu\left(S_{n}\right) \leq 3 n \epsilon+\epsilon
$$

and, since $\epsilon$ is arbitrary, $\mu\left(S_{n}\right)=0$ for any $n$. So, we have differentiability of $x \rightarrow \int_{x_{0}}^{x} f(y) d y$ almost everywhere.

Step 2. Next, we observe that if $f \in L_{1, \text { loc }}(\mathbb{R})$ satisfies

$$
\int_{\mathbb{R}} f \phi^{\prime} d x=0
$$

for any $\phi \in C_{0}^{\infty}(\mathbb{R})$, then $f=$ const almost everywhere. To prove this, we observe that if $\psi \in C_{0}^{\infty}(\mathbb{R})$ satisfies $\int_{\mathbb{R}} \psi d x=1$, then for any $\omega \in C_{0}^{\infty}(\mathbb{R})$ there is $\phi \in C_{0}^{\infty}(\mathbb{R})$ satisfying

$$
\phi^{\prime}=\omega-\psi \int_{\mathbb{R}} \omega d x
$$

Indeed, $h=\omega-\psi \int_{\mathbb{R}} \omega d x$ is continuous compactly supported with $\int_{\mathbb{R}} h d x=0$ and thus it has a unique compactly supported primitive.

Hence

$$
\int_{\mathbb{R}} f \phi^{\prime} d x=\int_{R} f\left(\omega-\psi \int_{\mathbb{R}} \omega d y\right) d x=0
$$

or

$$
\int_{\mathbb{R}}\left(f-\int_{\mathbb{R}} f \psi d y\right) \omega d x=0
$$

for any $\omega \in C_{0}^{\infty}(\mathbb{R})$ and thus $f=$ const almost everywhere.
Step 3. Next, if $v(x)=\int_{x_{0}}^{x} f(y) d y$ for $f \in L_{1, \text { loc }}(\mathbb{R})$, then $v$ is continuous and the generalized derivative of $v, D v$, equals $f$. In the proof, we can put $x_{0}=0$. Then

$$
\begin{aligned}
\int_{\mathbb{R}} v \phi^{\prime} d x & =\int_{0}^{\infty}\left(\int_{0}^{x} f(y) \phi^{\prime}(x) d y\right) d x-\int_{-\infty}^{0}\left(\int_{x}^{0} f(y) \phi^{\prime}(x) d y\right) d x \\
& =\int_{0}^{\infty} f(y)\left(\int_{y}^{\infty} \phi^{\prime}(x) d x\right) d y-\int_{-\infty}^{0} f(y)\left(\int_{-\infty}^{0} \phi^{\prime}(x) d x\right) d y \\
& =-\int_{\mathbb{R}} f(y) \phi(y) d y .
\end{aligned}
$$

With these results, let $u \in L_{1, \text { loc }}(\mathbb{R})$ be the distributional derivative $D u \in$ $L_{1, l o c}(\mathbb{R})$ and set $\bar{u}(x)=\int_{0}^{x} D u(t) d t$. Then $D \bar{u}=D u$ almost everywhere and hence $\bar{u}+C=u$ almost everywhere. Defining $\widetilde{u}=\bar{u}+C$, we see that $\widetilde{u}$ is continuous and has integral representation and thus it is differentiable almost everywhere.

### 1.2 Fundamental Theorems of Functional Analysis

The foundation of classical functional analysis are the four theorems which we formulate and discuss below.

### 1.2.1 Hahn-Banach Theorem

Theorem 1.12. (Hahn-Banach) Let $X$ be a normed space, $X_{0}$ a linear subspace of $X$, and $x_{1}^{*}$ a continuous linear functional defined on $X_{0}$. Then there exists a continuous linear functional $x^{*}$ defined on $X$ such that $x^{*}(x)=x_{1}^{*}(x)$ for $x \in X_{0}$ and $\left\|x^{*}\right\|=\left\|x_{1}^{*}\right\|$.

The Hahn-Banach theorem has a multitude of applications. For us, the most important one is in the theory of the dual space to $X$. The space $\mathcal{L}(X, \mathbb{R})$ (or $\mathcal{L}(X, \mathbb{C})$ ) of all continuous functionals is denoted by $X^{*}$ and referred to as the dual space. The Hahn-Banach theorem implies that $X^{*}$ is nonempty (as one can easily construct a continuous linear functional on a one-dimensional space) and, moreover, there are sufficiently many bounded functionals to separate points of $x$; that is, for any two points $x_{1}, x_{2} \in X$ there is $x^{*} \in X^{*}$ such that $x^{*}\left(x_{1}\right)=0$ and $x^{*}\left(x_{2}\right)=1$. The Banach space $X^{* *}=\left(X^{*}\right)^{*}$ is called the second dual. Every element $x \in X$ can be identified with an element of $X^{* *}$ by the evaluation formula

$$
\begin{equation*}
x\left(x^{*}\right)=x^{*}(x) ; \tag{1.20}
\end{equation*}
$$

that is, $X$ can be viewed as a subspace of $X^{* *}$. To indicate that there is some symmetry between $X$ and its dual and second dual we shall often write

$$
x^{*}(x)=<x^{*}, x>_{X^{*} \times X}
$$

where the subscript $X^{*} \times X$ is suppressed if no ambiguity is possible.
In general $X \neq X^{* *}$. Spaces for which $X=X^{* *}$ are called reflexive. Examples of reflexive spaces are rendered by Hilbert and $L_{p}$ spaces with $1<p<\infty$. However, the spaces $L_{1}$ and $L_{\infty}$, as well as nontrivial spaces of continuous functions, fail to be reflexive.

Example 1.13. If $1<p<\infty$, then the dual to $L_{p}(\Omega)$ can be identified with $L_{q}(\Omega)$ where $1 / p+1 / q=1$, and the duality pairing is given by

$$
\begin{equation*}
<f, g>=\int_{\Omega} f(\mathbf{x}) g(\mathbf{x}) d \mathbf{x}, \quad f \in L_{p}(\Omega), g \in L_{q}(\Omega) \tag{1.21}
\end{equation*}
$$

This shows, in particular, that $L_{2}(\Omega)$ is a Hilbert space and the above duality pairing gives the scalar product in the real case. If $L_{2}(\Omega)$ is considered over the complex field, then in order to get a scalar product, (1.21) should be modified by taking the complex adjoint of $g$.

Moreover, as mentioned above, the spaces $L_{p}(\Omega)$ with $1<p<\infty$ are reflexive. On the other hand, if $p=1$, then $\left(L_{1}(\Omega)\right)^{*}=L_{\infty}(\Omega)$ with duality pairing given again by (1.21). However, the dual to $L_{\infty}$ is much larger than $L_{1}(\Omega)$ and thus $L_{1}(\Omega)$ is not a reflexive space.

Another important corollary of the Hahn-Banach theorem is that for each $0 \neq x \in X$ there is an element $\bar{x}^{*} \in X^{*}$ that satisfies $\left\|\bar{x}^{*}\right\|=1$ and $<$ $\bar{x}^{*}, x>=\|x\|$. In general, the correspondence $x \rightarrow \bar{x}^{*}$ is multi-valued: this is the case in $L_{1}$-spaces and spaces of continuous functions it becomes, however, single-valued if the unit ball in $X$ is strictly convex (e.g., in Hilbert spaces or $L^{p}$-spaces with $1<p<\infty$; see [82]).

### 1.2.2 Spanning theorem and its application

A workhorse of analysis is the spanning criterion.
Theorem 1.14. Let $X$ be a normed space and $\left\{y_{j}\right\} \subset X$. Then $z \in Y:=$ $\overline{\mathcal{L}}$ in $\left\{y_{j}\right\}$ if and only if

$$
\forall_{x^{*} \in X^{*}}<x^{*}, y_{j}>=0 \quad \text { implies } \quad<x^{*}, z>=0
$$

Proof. In one direction it follows easily from linearity and continuity.
Conversely, assume $<x^{*}, z>=0$ for all $x^{*}$ annihilating $Y$ and $z \neq Y$. Thus, $\inf _{y \in Y}\|z-y\|=d>0$ (from closedness). Define $Z=\mathcal{L} i n\{Y, z\}$ and define a functional $y^{*}$ on $Z$ by $<y *, \xi>=<y^{*}, y+a z>=a$. We have

$$
\|y+a z\|=|a|\left\|\frac{y}{a}+z\right\| \geq|a| d
$$

hence

$$
\left|<y^{*}, \xi>=|a| \leq \frac{\|y+a z\|}{d}=d^{-1}\|\xi\|\right.
$$

and $y^{*}$ is bounded. By H.-B. theorem, we extend it to $\widetilde{y}^{*}$ on $X$ with $<\widetilde{y}^{*}, x>=$ 0 on $Y$ and $<\widetilde{y}^{*}, z>=1 \neq 0$.

Next we consider the Müntz theorem.
Theorem 1.15. Let $\left(\lambda_{j}\right)_{j \in \mathbb{N}}$ be a sequence of positive numbers tending to $\infty$. The functions $\left\{t^{\lambda_{j}}\right\}_{j \in \mathbb{N}}$ span the space of all continuous functions on $[0,1]$ that vanish at $t=0$ if and only if

$$
\sum_{j=1}^{\infty} \frac{1}{\lambda_{j}}=\infty
$$

Proof. We prove the 'sufficient' part. Let $x^{*}$ be a bounded linear functional that vanishes on all $t^{\lambda_{j}}$ :

$$
<x^{*}, t^{\lambda_{j}}>=0, \quad j \in \mathbb{N}
$$

For $\zeta \in \mathbb{C}$ such that $\Re \zeta>0$, the functions $\zeta \rightarrow t^{\zeta}$ are analytic functions with values in $C([0,1])$ This can be proved by showing that

$$
\lim _{\mathbb{C} \ni h \rightarrow 0} \frac{t^{\zeta+h}-t^{\zeta}}{h}=(\ln t) t^{\zeta}
$$

uniformly in $t \in[0,1]$. Then

$$
f(\zeta)=<x^{*}, t^{\zeta}>
$$

is a scalar analytic function of $\zeta$ with $\Re \zeta>0$. We can assume that $\left\|x^{*}\right\| \leq 1$. Then

$$
|f(\zeta)| \leq 1
$$

for $\Re \zeta>0$ and $f\left(\lambda_{j}\right)=0$ for any $j \in \mathbb{N}$.
Next, for a given $N$, we define a Blaschke product by

$$
B_{N}(\zeta)=\prod_{j=1}^{N} \frac{\zeta-\lambda_{j}}{\zeta+\lambda_{j}}
$$

We see that $B_{N}(\zeta)=0$ if and only if $\zeta=\lambda_{j},\left|B_{N}(\zeta)\right| \rightarrow 1$ both as $\Re \zeta \rightarrow 0$ and $|\zeta| \rightarrow \infty$. Hence

$$
g_{N}(\zeta)=\frac{f(\zeta)}{B_{N}(\zeta)}
$$

is analytic in $\Re \zeta>0$. Moreover, for any $\epsilon^{\prime}$ there is $\delta_{0}>0$ such that for any $\delta>\delta_{0}$ we have $\left|B_{N}(\zeta)\right| \geq 1-\epsilon^{\prime}$ on $\Re \zeta=\delta$ and $|\zeta|=\delta^{-1}$. Hence for any $\epsilon$

$$
\left|g_{N}(\zeta)\right| \leq 1+\epsilon
$$

there and by the maximum principle the inequality extends to the interior of the domain. Taking $\epsilon \rightarrow 0$ we obtain $\left|g_{N}(\zeta)\right| \leq 1$ on $\Re \zeta>0$.

Assume now there is $k>0$ for which $f(k) \neq 0$. Then we have

$$
\prod_{j=1}^{N}\left|\frac{\lambda_{j}+k}{\lambda_{j}-k}\right| \leq \frac{1}{f(k)}
$$

Note, that this estimate is uniform in $N$. If we write

$$
\frac{\lambda_{j}+k}{\lambda_{j}-k}=1+\frac{2 k}{\lambda_{j}-k}
$$

then, by $\lambda_{j} \rightarrow \infty$ almost all terms bigger then 1 . Remembering that boundedness of the product is equivalent to the boundedness of the sum

$$
\sum_{j=1}^{N} \frac{1}{\lambda_{j}-k}
$$

we see that we arrived at contradiction with the assumption. Hence, we must have $f(k)=0$ for any $k>0$. This means, however, that any functional that vanishes on $\left\{t^{\lambda_{j}}\right\}$ vanishes also on $t^{k}$ for any $k$. But, by the Stone- Weierstrass theorem, it must vanish on any continuous function (taking value 0 at zero). Hence, by the spanning criterion, any such continuous function belongs to the closed linear span of $\left\{t^{\lambda_{j}}\right\}$.

Example 1.16. The existence of an element $\bar{x}^{*}$ satisfying $<\bar{x}^{*}, x>=\|x\|$ has an important consequence for the relation between $X$ and $X^{* *}$ in a nonreflexive case. Let $B, B^{*}, B^{* *}$ denote the unit balls in $X, X^{*}, X^{* *}$, respectively. Because $x^{*} \in X^{*}$ is an operator over $X$, the definition of the operator norm gives

$$
\begin{equation*}
\left\|x^{*}\right\|_{X^{*}}=\sup _{x \in B}\left|<x^{*}, x>\right|=\sup _{x \in B}<x^{*}, x>, \tag{1.22}
\end{equation*}
$$

and similarly, for $x \in X$ considered as an element of $X^{* *}$ according to (1.20), we have

$$
\begin{equation*}
\|x\|_{X^{* *}}=\sup _{x^{*} \in B^{*}}\left|<x^{*}, x>\right|=\sup _{x^{*} \in B^{*}}<x^{*}, x> \tag{1.23}
\end{equation*}
$$

Thus, $\|x\|_{X^{* *}} \leq\|x\|_{X}$. On the other hand,

$$
\|x\|_{X}=<\bar{x}^{*}, x>\leq \sup _{x^{*} \in B^{*}}<x^{*}, x>=\|x\|_{X^{* *}}
$$

and

$$
\begin{equation*}
\|x\|_{X^{* *}}=\|x\|_{X} \tag{1.24}
\end{equation*}
$$

Hence, in particular, the identification given by (1.20) is an isometry and $X$ is a closed subspace of $X^{* *}$.

The existence of a large number of functionals over $X$ allows us to introduce new types of convergence. Apart from the standard norm (or strong) convergence where $\left(x_{n}\right)_{n \in \mathbb{N}} \subset X$ converges to $x$ if

$$
\lim _{n \rightarrow \infty}\left\|x_{n}-x\right\|=0
$$

we define weak convergence by saying that $\left(x_{n}\right)_{n \in \mathbb{N}}$ weakly converges to $x$, if for any $x^{*} \in X^{*}$,

$$
\lim _{n \rightarrow \infty}<x^{*}, x_{n}>=<x^{*}, x>
$$

In a similar manner, we say that $\left(x_{n}^{*}\right)_{n \in \mathbb{N}} \subset X^{*}$ converges $*$-weakly to $x^{*}$ if, for any $x \in X$,

$$
\lim _{n \rightarrow \infty}<x_{n}^{*}, x>=<x^{*}, x>
$$

Remark 1.17. It is worthwhile to note that we have a concept of a weakly convergent or weakly Cauchy sequence if the finite limit $\lim _{n \rightarrow \infty}<x^{*}, x_{n}>$ exists for any $x^{*} \in X^{*}$. In general, in this case we do not have a limit element. If every weakly convergent sequence converges weakly to an element of $X$, the Banach space is said to be weakly sequentially complete. It can be proved that reflexive spaces and $L_{1}$ spaces are weakly sequentially complete. On the other hand, no space containing a subspace isomorphic to the space $c_{0}$ (of sequences that converge to 0 ) is weakly sequentially complete (see, e.g., [6]).

Remark 1.18. Weak convergence is indeed weaker than the convergence in norm. However, we point out that a theorem proved by Mazur (e.g., see [172], p. 120) says that if $x_{n} \rightarrow x$ weakly, then there is a sequence of convex combinations of elements of $\left(x_{n}\right)_{n \in \mathbb{N}}$ that converges to $x$ in norm. Thus, in particular, the norm and the weak closure of a convex sets coincide.

### 1.2.3 Banach-Steinhaus Theorem

Another fundamental theorem of functional analysis is the Banach-Steinhaus theorem, or the Uniform Boundedness Principle. It is based on a fundamental topological results known as the Baire Category Theorem.
Theorem 1.19. Let $X$ be a complete metric space and let $\left\{X_{n}\right\}_{n \geq 1}$ be a sequence of closed subsets in $X$. If Int $X_{n}=\emptyset$ for any $n \geq 1$, then Int $\bigcup_{n=1}^{\infty} X_{n}=\emptyset$. Equivalently, taking complements, we can state that a countable intersection of open dense sets is dense.

Remark 1.20. Baire's theorem is often used in the following equivalent form: if $X$ is a complete metric space and $\left\{X_{n}\right\}_{n \geq 1}$ is a countable family of closed sets such that $\bigcup_{n=1}^{\infty} X_{n}=X$, then $\operatorname{Int} X_{n} \neq \emptyset$ at least for one $n$.

## Chaotic dynamical systems

We assume that $X$ is a complete metric space, called the state space. In general, a dynamical system on $X$ is just a family of states $(\mathbf{x}(t))_{t \in \mathbb{T}}$ parametrized
by some parameter $t$ (time). Two main types of dynamical systems occur in applications: those for which the time variable is discrete (like the observation times) and those for which it is continuous.

Theories for discrete and continuous dynamical systems are to some extent parallel. In what follows mainly we will be concerned with continuous dynamical systems. Also, to fix attention we shall discuss only systems defined for $t \geq 0$, that are sometimes called semidynamical systems. Thus by a continuous dynamical system we will understand a family of functions (operators) $(\mathbf{x}(t, \cdot))_{t \geq 0}$ such that for each $t, \mathbf{x}(t, \cdot): X \rightarrow X$ is a continuous function, for each $\mathbf{x}_{0}$ the function $t \rightarrow \mathbf{x}\left(t, \mathbf{x}_{0}\right)$ is continuous with $\mathbf{x}\left(0, \mathbf{x}_{0}\right)=\mathbf{x}_{0}$. Moreover, typically it is required that the following semigroup property is satisfied (both in discrete and continuous case)

$$
\begin{equation*}
\mathbf{x}\left(t+s, \mathbf{x}_{0}\right)=\mathbf{x}\left(t, \mathbf{x}\left(s, \mathbf{x}_{0}\right)\right), \quad t, s \geq 0 \tag{1.25}
\end{equation*}
$$

which expresses the fact that the final state of the system can be obtained as the superposition of intermediate states.

Often discrete dynamical systems arise from iterations of a function

$$
\begin{equation*}
\mathbf{x}\left(t+1, \mathbf{x}_{0}\right)=f\left(\mathbf{x}\left(t, \mathbf{x}_{0}\right)\right), \quad t \in \mathbb{N} \tag{1.26}
\end{equation*}
$$

while when $t$ is continuous, the dynamics are usually described by a differential equation

$$
\begin{equation*}
\frac{d \mathbf{x}}{d t}=\dot{\mathbf{x}}=A(\mathbf{x}), \quad \mathbf{x}(0)=\mathbf{x}_{0} \quad t \in \mathbb{R}_{+} \tag{1.27}
\end{equation*}
$$

Let $(X, d)$ be a metric space where, to avoid non-degeneracy, we assume that $X \neq\{\mathbf{x}(t, \boldsymbol{p})\}_{t \geq 0}$ for any $\boldsymbol{p} \in X$, that is, the space does not degenerates to a single orbit). We say that the dynamical system $(\mathbf{x}(t))_{t \geq 0}$ on $(X, d)$ is topologically transitive if for any two non-empty open sets $U, \bar{V} \subset X$ there is $t_{0} \geq 0$ such that $\mathbf{x}(t, U) \cap V \neq \emptyset$. A periodic point of $(\mathbf{x}(t))_{t \geq 0}$ is any point $\boldsymbol{p} \in X$ satisfying

$$
\mathbf{x}(T, \boldsymbol{p})=\boldsymbol{p}
$$

for some $T>0$. The smallest such $T$ is called the period of $\boldsymbol{p}$. We say that the system has sensitive dependence on initial conditions, abbreviated as sdic, if there exists $\delta>0$ such that for every $\boldsymbol{p} \in X$ and a neighbourhood $N_{p}$ of $\boldsymbol{p}$ there exists a point $\mathbf{y} \in N_{p}$ and $t_{0}>0$ such that the distance between $\mathbf{x}\left(t_{0}, \boldsymbol{p}\right)$ and $\mathbf{x}\left(t_{0}, \mathbf{y}\right)$ is larger than $\delta$. This property captures the idea that in chaotic systems minute errors in experimental readings eventually lead to large scale divergence, and is widely understood to be the central idea in chaos.

With this preliminaries we are able to state Devaney's definition of chaos (as applied to continuous dynamical systems).

Definition 1.21. Let $X$ be a metric space. A dynamical system $(\mathbf{x}(t))_{t \geq 0}$ in $X$ is said to be chaotic in $X$ if

1. $(\mathbf{x}(t))_{t \geq 0}$ is transitive,
2. the set of periodic points of $(\mathbf{x}(t))_{t \geq 0}$ is dense in $X$,
3. $(\mathbf{x}(t))_{t \geq 0}$ has sdic.

To summarize, chaotic systems have three ingredients: indecomposability (property 1), unpredictability (property 3 ), and an element of regularity (property 2).

It is then a remarkable observation that properties 1 . and 2 together imply sdic.

Theorem 1.22. If $(\mathbf{x}(t))_{t \geq 0}$ is topologically transitive and has dense set of periodic points, then it has sdic.

We say that $X$ is non-degenerate, if continuous images of a compact intervals are nowhere dense in $X$.

Lemma 1.23. Let $X$ be a non-degenerate metric space. If the orbit $O(\boldsymbol{p})=$ $\{\mathbf{x}(t, \boldsymbol{p})\}_{t \geq 0}$ is dense in $X$, then also the orbit $O(\mathbf{x}(s, \boldsymbol{p}))=\{\mathbf{x}(t, \boldsymbol{p})\}_{t>s}$ is dense in $X$, for any $s>0$.

Proof. Assume that $O(\mathbf{x}(s, \boldsymbol{p}))$ is not dense in $X$, then there is an open ball $B$ such that $B \cap \overline{O(\mathbf{x}(s, \boldsymbol{p}))}=\emptyset$. However, each point of the ball is a limit point of the whole orbit $O(\boldsymbol{p})$, thus we must have $\{\mathbf{x}(t, \boldsymbol{p})\}_{0 \leq t \leq s}=$ $\overline{\{\mathbf{x}(t, \boldsymbol{p})\}_{0 \leq t \leq s}} \supset B$ which contradicts the assumption of nondegeneracy.

To fix terminology we say that a semigroup having a dense trajectory is called hypercyclic. We note that by continuity $\overline{O(\boldsymbol{p})}=\overline{\{\mathbf{x}(t, \boldsymbol{p})\}_{t \in \mathbb{Q}}}$, where $\mathbb{Q}$ is the set of positive rational numbers, therefore hypercyclic semigroups can exist only in separable spaces.

By $X_{h}$ we denote the set of hypercyclic vectors, that is,

$$
X_{h}=\{\boldsymbol{p} \in X ; O(\boldsymbol{p}) \text { is dense in } X\}
$$

Note that if $(\mathbf{x}(t))_{t \geq 0}$ has one hypercyclic vector, then it has a dense set of hypercyclic vectors as each of the point on the orbit $O(\boldsymbol{p})$ is hypercyclic (by the first part of the proof above).

Theorem 1.24. Let $(\mathbf{x}(t))_{t \geq 0}$ be a strongly continuous semigroup of continuous operators (possibly nonlinear) on a complete (separable) metric space $X$. The following conditions are equivalent:

1. $X_{h}$ is dense in $X$,
2. $(\mathrm{x}(t))_{t \geq 0}$ is topologically transitive.

Proof. Let as take the set of nonegative rational numbers and enumerate them as $\left\{t_{1}, t_{2}, \ldots\right\}$. Consider now the family $\left\{\mathbf{x}\left(t_{n}\right)\right\}_{n \in \mathbb{N}}$. Clearly, the orbit of $\boldsymbol{p}$ through $(\mathbf{x}(t))_{t \geq 0}$ is dense in $X$ if and only if the set $\left\{\mathbf{x}\left(t_{n}\right) \boldsymbol{p}\right\}_{n \in \mathbb{N}}$ is dense.

Consider now the covering of $X$ by the enumerated sequence of balls $B_{m}$ centered at points of a countable subset of $X$ with rational radii. Since each $\mathbf{x}\left(t_{m}\right)$ is continuous, the sets

$$
G_{m}=\bigcup_{n \in \mathbb{N}} \mathbf{x}^{-1}\left(t_{n}, B_{m}\right)
$$

are open. Next we claim that

$$
X_{h}=\bigcap_{m \in \mathbb{N}} G_{m}
$$

In fact, let $\boldsymbol{p} \in X_{h}$, that is, $\boldsymbol{p}$ is hypercyclic. It means that $\mathbf{x}\left(t_{n}, \boldsymbol{p}\right)$ visits each neigbourhood of each point of $X$ for some $n$. In particular, for each $m$ there must be $n$ such that $\mathbf{x}\left(t_{n}, \boldsymbol{p}\right) \in B_{m}$ or $\boldsymbol{p} \in \mathbf{x}^{-1}\left(t_{n}, B_{m}\right)$ which means $\boldsymbol{p} \in \bigcap_{m \in \mathbb{N}} G_{m}$.

Conversely, if $\boldsymbol{p} \in \bigcap_{m \in \mathbb{N}} G_{m}$, then for each $m$ there is $n$ such that $\boldsymbol{p} \in$ $\mathbf{x}^{-1}\left(t_{n}, B_{m}\right)$, that is, $\mathbf{x}\left(t_{n}, \boldsymbol{p}\right) \in B_{m}$. This means that $\left\{\mathbf{x}\left(t_{n}, \boldsymbol{p}\right)\right\}_{n \in \mathbb{N}}$ is dense.

The next claim is condition 2 . is equivalent to each set $G_{m}$ being dense in $X$. If $G_{m}$ were not dense, then for some $B_{r}, B_{r} \cap \mathbf{x}^{-1}\left(t_{n}, B_{m}\right)=\emptyset$ for any $n$. But then $\mathbf{x}\left(t_{n}, B_{r}\right) \cap B_{m}=\emptyset$ for any $n$. Since the continuous semigroup is topologically transitive, we know that there is $\mathbf{y} \in B_{r}$ such that $\mathbf{x}\left(t_{0}, \mathbf{y}\right) \in B_{m}$ for some $t_{0}$. Since $B_{m}$ is open, $\mathbf{x}(t, \mathbf{y}) \in B_{m}$ for $t$ from some neighbourhood of $t_{0}$ and this neighbourhood must contain rational numbers.

The converse is immediate as for given open $U$ and $V$ we find $B_{m} \subset V$ and since $G_{m}$ is dense $U \cap G_{m} \neq \emptyset$. Thus $U \cap \mathbf{x}^{-1}\left(t_{n}, B_{m}\right) \neq \emptyset$ for some $n$, hence $\mathbf{x}\left(t_{n}, U\right) \cap B_{m} \neq \emptyset$.

So, if $(\mathbf{x}(t))_{t \geq 0}$ is topologically transitive, then $X_{h}$ is the intersection of a countable collection of open dense sets, and by Baire Theorem in a complete space such an intersection must be still dense, thus $X_{h}$ is dense.

Conversely, if $X_{h}$ is dense, then each term of the intersection must be dense, thus each $G_{m}$ is dense which yields the transitivity.

## Back to Banach-Steinhaus Theorem

To understand its importance, let us reflect for a moment on possible types of convergence of sequences of operators. Because the space $\mathcal{L}(X, Y)$ can be made a normed space by introducing the norm (1.11), the most natural concept of convergence of a sequence $\left(A_{n}\right)_{n \in \mathbb{N}}$ would be with respect to this norm. Such a convergence is referred to as the uniform operator convergence. However, for many purposes this notion is too strong and we work with the pointwise or strong convergence: the sequence $\left(A_{n}\right)_{n \in \mathbb{N}}$ is said to converge strongly if, for each $x \in X$, the sequence $\left(A_{n} x\right)_{n \in \mathbb{N}}$ converges in the norm of $Y$. In the same way we define uniform and strong boundedness of a subset of $\mathcal{L}(X, Y)$.

Note that if $Y=\mathbb{R}($ or $\mathbb{C})$, then strong convergence coincides with $*$-weak convergence.

After these preliminaries we can formulate the Banach-Steinhaus theorem.

Theorem 1.25. Assume that $X$ is a Banach space and $Y$ is a normed space. Then a subset of $\mathcal{L}(X, Y)$ is uniformly bounded if and only if it is strongly bounded.

One of the most important consequences of the Banach-Steinhaus theorem is that a strongly converging sequence of bounded operators is always converging to a linear bounded operator. That is, if for each $x$ there is $y_{x}$ such that

$$
\lim _{n \rightarrow \infty} A_{n} x=y_{x}
$$

then there is $A \in \mathcal{L}(X, Y)$ satisfying $A x=y_{x}$.
Example 1.26. We can use the above result to get a better understanding of the concept of weak convergence and, in particular, to clarify the relation between reflexive and weakly sequentially complete spaces. First, by considering elements of $X^{*}$ as operators in $\mathcal{L}(X, \mathbb{C})$, we see that every $*$-weakly converging sequence of functionals converges to an element of $X^{*}$ in $*$-weak topology. On the other hand, for a weakly converging sequence $\left(x_{n}\right)_{n \in \mathbb{N}} \subset X$, such an approach requires that $x_{n}, n \in \mathbb{N}$, be identified with elements of $X^{* *}$ and thus, by the Banach-Steinhaus theorem, a weakly converging sequence always has a limit $x \in X^{* *}$. If $X$ is reflexive, then $x \in X$ and $X$ is weakly sequentially complete. However, for nonreflexive $X$ we might have $x \in X^{* *} \backslash X$ and then $\left(x_{n}\right)_{n \in \mathbb{N}}$ does not converge weakly to any element of $X$.

On the other hand, (1.24) implies that a weakly convergent sequence is norm bounded.

We note another important corollary of the Banach-Steinhaus theorem which we use in the sequel.

Corollary 1.27. A sequence of operators $\left(A_{n}\right)_{n \in \mathbb{N}}$ is strongly convergent if and only if it is convergent uniformly on compact sets.

Proof. It is enough to consider convergence to 0 . If $\left(A_{n}\right)_{n \in \mathbb{N}}$ converges strongly, then by the Banach-Steinhaus theorem, $a=\sup _{n \in \mathbb{N}}\left\|A_{n}\right\|<+\infty$. Next, if $\Omega \subset X$ is compact, then for any $\epsilon$ we can find a finite set $N_{\epsilon}=$ $\left\{x_{1}, \ldots, x_{k}\right\}$ such that for any $x \in \Omega$ there is $x_{i} \in N_{\epsilon}$ with $\left\|x-x_{i}\right\| \leq \epsilon / 2 a$. Because $N_{\epsilon}$ is finite, we can find $n_{0}$ such that for all $n>n_{0}$ and $i=1, \ldots, k$ we have $\left\|A_{n} x_{i}\right\| \leq \epsilon / 2$ and hence

$$
\left\|A_{n} x\right\|=\left\|A_{n} x_{i}\right\|+a\left\|x-x_{i}\right\| \leq \epsilon
$$

for any $x \in \Omega$. The converse statement is obvious.
We conclude this unit by presenting a frequently used result related to the Banach-Steinhaus theorem.
Proposition 1.28. Let $X, Y$ be Banach spaces and $\left(A_{n}\right)_{n \in \mathbb{N}} \subset \mathcal{L}(X, Y)$ be a sequence of operators satisfying $\sup _{n \in \mathbb{N}}\left\|A_{n}\right\| \leq M$ for some $M>0$. If there is a dense subset $D \subset X$ such that $\left(A_{n} x\right)_{n \in \mathbb{N}}$ is a Cauchy sequence for any $x \in D$, then $\left(A_{n} x\right)_{n \in \mathbb{N}}$ converges for any $x \in X$ to some $A \in \mathcal{L}(X, Y)$.

Proof. Let us fix $\epsilon>0$ and $y \in X$. For this $\epsilon$ we find $x \in D$ with $\|x-y\|<\epsilon / M$ and for this $x$ we find $n_{0}$ such that $\left\|A_{n} x-A_{m} x\right\|<\epsilon$ for all $n, m>n_{0}$. Thus,

$$
\left\|A_{n} y-A_{m} y\right\| \leq\left\|A_{n} x-A_{m} x\right\|+\left\|A_{n}(x-y)\right\|+\left\|A_{m}(x-y)\right\| \leq 3 \epsilon
$$

Hence, $\left(A_{n} y\right)_{n \in \mathbb{N}}$ is a Cauchy sequence for any $y \in X$ and, because $Y$ is a Banach space, it converges and an application of the Banach-Steinhaus theorem ends the proof.

## Application-limits of integral expressions

Consider an equation describing growth of, say, cells

$$
\begin{equation*}
\frac{\partial N}{\partial t}+\frac{\partial(g(m) N)}{\partial m}=-\mu(m) N(t, m), \quad m \in(0,1) \tag{1.28}
\end{equation*}
$$

with the boundary condition

$$
\begin{equation*}
g(0) N(t, 0)=0 \tag{1.29}
\end{equation*}
$$

and with the initial condition

$$
\begin{equation*}
N(0, m)=N_{0}(m) \quad \text { for } m \in[0,1] \tag{1.30}
\end{equation*}
$$

Here $N(m)$ denotes cells' density with respect to their size/mass and we consider the problem in $L_{1}([0,1])$.

Consider the 'formal' equation for the stationary version of the equation (the resolvent equation)

$$
\lambda N(m)+(g(m) N(m))^{\prime}+\mu(m) N(m)=f(m) \in L_{1}([0,1])
$$

whose solution is given by

$$
\begin{equation*}
N_{\lambda}(m)=\frac{e^{-\lambda G(m)-Q(m)}}{g(m)} \int_{0}^{m} e^{\lambda G(s)+Q(s)} f(s) d s \tag{1.31}
\end{equation*}
$$

where $G(m)=\int_{0}^{m}(1 / g(s)) d s$ and $Q(m)=\int_{0}^{m}(\mu(s) / g(s)) d s$. To shorten notation we denote

$$
e_{-\lambda}(m):=e^{-\lambda G(m)-Q(m)}, \quad e_{\lambda}(m):=e^{\lambda G(m)+Q(m)}
$$

Our aim is to show that $g(m) N_{\lambda}(m) \rightarrow 0$ as $m \rightarrow 1^{-}$provided $1 / g$ or $\mu$ is not integrable close to 1 . If the latter condition is satisfied, then $e_{\lambda}(m) \rightarrow \infty$ and $e^{-\lambda}(m) \rightarrow 0$ as $m \rightarrow 1^{-}$.

Indeed, consider the family of functionals $\left\{\xi_{m}\right\}_{m \in[1-\epsilon, 1)}$ for some $\epsilon>0$ defined by

$$
\xi_{m} f=e_{-\lambda}(m) \int_{0}^{m} e_{\lambda}(s) f(s) d s
$$

for $f \in L^{1}[0,1]$. We have

$$
\left|\xi_{m} f\right| \leq e_{-\lambda}(m) \int_{0}^{m} e_{\lambda}(s)|f(s)| d s \leq \int_{0}^{1}|f(s)| d s
$$

on account of monotonicity of $e_{\lambda}$. Moreover, for $f$ with support in $[0,1-\delta]$ with any $\delta>0$ we have $\lim _{m \rightarrow 1^{-}} \xi_{m} f=0$ and, by Proposition 1.28 , the above limit extends by density for any $f \in L^{1}[0,1]$.

### 1.2.4 Weak compactness

In finite dimensional spaces normed spaces we have Bolzano-Weierstrass theorem stating that from any bounded sequence of elements of $X_{n}$ one can select a convergent subsequence. In other words, a closed unit ball in $X_{n}$ is compact.

There is no infinite dimensional normed space in which the unit ball is compact.

Weak compactness comes to the rescue. Let us begin with (separable) Hilbert spaces.

Theorem 1.29. Each bounded sequence $\left(u_{n}\right)_{n \in \mathbb{N}}$ in a separable Hilbert space $X$ has a weakly convergent subsequence.

Proof. Let $\left\{v_{k}\right\}_{k \in \mathbb{N}}$ be dense in $X$ and consider numerical sequences $\left(\left(u_{n}, v_{k}\right)\right)_{n \in \mathbb{N}}$ for any $k$. From Banach-Steinhaus theorem and

$$
\left|\left(u_{n}, v_{k}\right)\right| \leq\left\|u_{n}\right\|\left\|v_{k}\right\|
$$

we see that for each $k$ these sequences are bounded and hence each has a convergent subsequence. We use the diagonal procedure: first we select $\left(u_{1 n}\right)_{n \in \mathbb{N}}$ such that $\left(u_{1 n}, v_{1}\right) \rightarrow a_{1}$, then from $\left(u_{1 n}\right)_{n \in \mathbb{N}}$ we select $\left(u_{2 n}\right)_{n \in \mathbb{N}}$ such that $\left(u_{2 n}, v_{2}\right) \rightarrow a_{2}$ and continue by induction. Finally, we take the diagonal sequence $w_{n}=u_{n n}$ which has the property that $\left(w_{n}, v_{k}\right) \rightarrow a_{k}$. This follows from the fact that elements of $\left(w_{n}\right)_{n \in \mathbb{N}}$ belong to ( $u_{k n}$ for $n \geq k$. Since $\left\{v_{k}\right\}_{k \in \mathbb{N}}$ is dense in $X$ and $\left(u_{n}\right)_{n \in \mathbb{N}}$ is norm bounded, Proposition 1.28 implies $\left(\left(w_{n}, v\right)\right)_{n \in \mathbb{N}}$ converges to, say, $a(v)$ for any $v \in X$ and $v \rightarrow a(v)$ is a bounded (anti) linear functional on $X$. By the Riesz representation theorem, there is $w \in X$ such that $a(v)=(v, w)$ and thus $w_{n} \rightharpoonup w$.

If $X$ is not separable, then we can consider $Y=\overline{\mathcal{L} i n\left\{u_{n}\right\}_{n \in \mathbb{N}}}$ which is separable and apply the above theorem in $Y$ getting an element $w \in Y$ for which

$$
\left(w_{n}, v\right) \rightarrow(w, v), \quad v \in Y
$$

Let now $z \in X$. By orthogonal decomposition, $z=v+v^{\perp}$ by linearity and continuity (as $w \in Y$ )

$$
\left(w_{n}, z\right)=\left(w_{n}, v\right) \rightarrow(w, v)=(w, z)
$$

and so $w_{n} \rightharpoonup w$ in $X$.

Corollary 1.30. Closed unit ball in $X$ is weakly sequentially compact.
Proof. We have

$$
\left(v, w_{n}\right) \rightarrow(v, w), \quad n \rightarrow \infty
$$

for any $v$. We can assume $w=0$ We prove that for any $k$ there are indices $n_{1}, \ldots, n_{k}$ such that

$$
k^{-1}\left(w_{n_{1}}+\ldots+w_{n_{k}}\right) \rightarrow 0
$$

in $X$. Since $\left(w_{1}, w_{n}\right) \rightarrow 0$, we set $n_{1}=1$ and select $n_{2}$ such that $\left|\left(w_{n_{1}}, w_{n_{2}}\right)\right| \leq$ $1 / 2$. Then we select $n_{3}$ such that $\left|\left(w_{n_{1}}, w_{n_{3}}\right)\right| \leq 1 / 2$ and $\left|\left(w_{n_{2}}, w_{n_{3}}\right)\right| \leq 1 / 2$ and further, $n_{k}$ such that $\left|\left(w_{n_{1}}, w_{n_{k}}\right)\right| \leq 1 /(k-1), \ldots,\left|\left(w_{n_{k-1}}, w_{n_{k}}\right)\right| \leq 1 /(k-$ 1). Since $\left\|w_{n}\right\| \leq C$, we obtain

$$
\begin{aligned}
& \left\|k^{-1}\left(w_{n_{1}}+\ldots+w_{n_{k}}\right)\right\|^{2} \\
& \leq k^{-2}\left(\sum_{j=1}^{k}\left\|w_{n_{j}}\right\|^{2}+2 \sum_{j=1}^{k-1}\left(w_{n_{j}}, w_{n_{k}}\right)+2 \sum_{j=1}^{k-2}\left(w_{n_{j}}, w_{n_{k-1}}\right)+\ldots\right) \\
& \leq k^{-2}\left(k C^{2}+2(k-1)(k-1)^{-1}+2(k-2)(k-2)^{-1}+\ldots 2\right) \\
& \leq k^{-1}\left(C^{2}+2\right)
\end{aligned}
$$

Note that this result shows that any closed convex set in $X$ is weakly sequentially compact. What about other spaces?

Practically the same proof (using the fact that a closed subspace of a reflexive space is reflexive) shows that if a Banach space is reflexive, then the closed unit ball is weakly sequentially compact. The converse is also true (Eberlain).

Helly's theorem: If $X$ is a separable Banach space and $U=X^{*}$, then the closed unit ball in $U$ is weak* sequentially compact. Alaoglu removed separability.

### 1.2.5 The Open Mapping Theorem

The Open Mapping Theorem is fundamental for inverting linear operators. Let us recall that an operator $A: X \rightarrow Y$ is called surjective if $\operatorname{Im} A=Y$ and open if the set $A \Omega$ is open for any open set $\Omega \subset X$.

Theorem 1.31. Let $X, Y$ be Banach spaces. Any surjective $A \in \mathcal{L}(X, Y)$ is an open mapping.

One of the most often used consequences of this theorem is the Bounded Inverse Theorem.

Corollary 1.32. If $A \in \mathcal{L}(X, Y)$ is such that $\operatorname{Ker} A=\{0\}$ and $\operatorname{Im} A=Y$, then $A^{-1} \in \mathcal{L}(Y, X)$.

The corollary follows as the assumptions on the kernel and the image ensure the existence of a linear operator $A^{-1}$ defined on the whole $Y$. The operator $A^{-1}$ is continuous by the Open Mapping Theorem, as the preimage of any open set in $X$ through $A^{-1}$, that is, the image of this set through $A$, is open.

Throughout the book we are faced with invertibility of unbounded operators. An operator $(A, D(A))$ is said to be invertible if there is a bounded operator $A^{-1} \in \mathcal{L}(Y, X)$ such that $A^{-1} A x=x$ for all $x \in D(A)$ and $A^{-1} y \in D(A)$ with $A A^{-1} y=y$ for any $y \in Y$. We have the following useful conditions for invertibility of $A$.

Proposition 1.33. Let $X, Y$ be Banach spaces and $A \in L(X, Y)$. The following assertions are equivalent.
(i) $A$ is invertible;
(ii) $\operatorname{Im} A=Y$ and there is $m>0$ such that $\|A x\| \geq m\|x\|$ for all $x \in D(A)$;
(iii) $A$ is closed, $\overline{\operatorname{ImA}}=Y$ and there is $m>0$ such that $\|A x\| \geq m\|x\|$ for all $x \in D(A)$;
(iv) $A$ is closed, $\operatorname{Im} A=Y$, and $\operatorname{Ker} A=\{0\}$.

Proof. The equivalence of (i) and (ii) follows directly from the definition of invertibility. By Theorem 1.34, the graph of any bounded operator is closed and because the graph of the inverse is given by

$$
G(A)=\left\{(x, y) ; \quad(y, x) \in G\left(A^{-1}\right)\right\}
$$

we see that the graph of any invertible operator is closed and thus any such an operator is closed. Hence, (i) and (ii) imply (iii) and (iv). Assume now that (iii) holds. $G(A)$ is a closed subspace of $X \times Y$, therefore it is a Banach space itself. The inequality $\|A x\| \geq m\|x\|$ implies that the mapping $G(A) \ni$ $(x, A x) \rightarrow A x \in \operatorname{Im} A$ is an isomorphism onto $\operatorname{Im} A$ and hence $\operatorname{Im} A$ is also closed. Thus $\operatorname{Im} A=Y$ and (ii) follows. Finally, if (iv) holds, then Corollary 1.32 can be applied to $A$ from $D(A)$ (with the graph norm) to $Y$ to show that $A^{-1} \in \mathcal{L}(Y, D(A)) \subset \mathcal{L}(Y, X)$.

Norm equivalence. An important result is that if $X$ is a Banach space with respect to two norms $\|\cdot\|_{1}$ and $\|\cdot\|_{2}$ and there is $C$ such that $\|x\|_{1} \leq C\|x\|_{2}$, then both norms are equivalent.

## The Closed Graph Theorem

It is easy to see that a bounded operator defined on the whole Banach space $X$ is closed. That the inverse also is true follows from the Closed Graph Theorem.

Theorem 1.34. Let $X, Y$ be Banach spaces. An operator $A \in L(X, Y)$ with $D(A)=X$ is bounded if and only if its graph is closed.

We can rephrase this result by saying that an everywhere defined closed operator in a Banach space must be bounded.

Proof. Indeed, consider on $X$ two norms, the original norm $\|\cdot\|$ and the graph norm

$$
\|x\|_{D(A)}=\|x\|+\|A x\|
$$

By closedness, $X$ is a Banach space with respect to $D(A)$ and $A$ is continuous in the norm $\|\cdot\|_{D(A)}$. Hence, the norms are equivalent and $A$ is continuous in the norm $\|\cdot\|$.

To give a nice and useful example of an application of the Closed Graph Theorem, we discuss a frequently used notion of relatively bounded operators. Let two operators $(A, D(A))$ and $(B, D(B))$ be given. We say that $B$ is $A$ bounded if $D(A) \subset D(B)$ and there exist constants $a, b \geq 0$ such that for any $x \in D(A)$,

$$
\begin{equation*}
\|B x\| \leq a\|A x\|+b\|x\| \tag{1.32}
\end{equation*}
$$

Note that the right-hand side defines a norm on the space $D(A)$, which is equivalent to the graph norm (1.15).

Corollary 1.35. If $A$ is closed and $B$ closable, then $D(A) \subset D(B)$ implies that $B$ is $A$-bounded.

Proof. If $A$ is a closed operator, then $D(A)$ equipped with the graph norm is a Banach space. If we assume that $D(A) \subset D(B)$ and $(B, D(B))$ is closable, then $D(A) \subset D(\bar{B})$. Because the graph norm on $D(A)$ is stronger than the norm induced from $X$, the operator $\bar{B}$, considered as an operator from $D(A)$ to $X$ is everywhere defined and closed. On the other hand, $\left.\bar{B}\right|_{D(A)}=B$; hence $B: D(A) \rightarrow X$ is bounded by the Closed Graph Theorem and thus $B$ is $A$-bounded.

### 1.3 Hilbert space methods

One of the most often used theorems of functional analysis is the Riesz representation theorem.

Theorem 1.36 (Riesz representation theorem). If $x^{*}$ is a continuous linear functional on a Hilbert space $H$, then there is exactly one element $y \in H$ such that

$$
\begin{equation*}
<x^{*}, x>=(x, y) \tag{1.33}
\end{equation*}
$$

### 1.3.1 To identify or not to identify-the Gelfand triple

Riesz theorem shows that there is a canonical isometry between a Hilbert space $H$ and its dual $H^{*}$. It is therefore natural to identify $H$ and $H^{*}$ and is
done so in most applications. There are, however, situations when it cannot be done.

Assume that $H$ is a Hilbert space equipped with a scalar product $(\cdot, \cdot)_{H}$ and that $V \subset H$ is a subspace of $H$ which is a Hilbert space in its own right, endowed with a scalar product $(\cdot, \cdot)_{V}$. Assume that $V$ is densely and continuously embedded in $H$ that is $\bar{V}=H$ and $\|x\|_{H} \leq c\|x\|_{V}, x \in V$, for some constant $c$. There is a canonical map $T: H^{*} \rightarrow V^{*}$ which is given by restriction to $V$ of any $h^{*} \in H^{*}$ :

$$
<T h^{*}, v>_{V^{*} \times V}=<h^{*}, v>_{H^{*} \times H}, \quad v \in V
$$

We easily see that

$$
\left\|T h^{*}\right\|_{V^{*}} \leq C\left\|h^{*}\right\|_{H^{*}}
$$

Indeed

$$
\begin{aligned}
\left\|T h^{*}\right\|_{V^{*}} & =\sup _{\|v\|_{V} \leq 1}\left|<T h^{*}, v>_{V^{*} \times V}\right|=\sup _{\|v\|_{V} \leq 1}\left|<h^{*}, v>_{H^{*} \times H}\right| \\
& \leq\left\|h^{*}\right\|_{H^{*}} \sup _{\|v\|_{V} \leq 1}\|v\|_{H} \leq c\left\|h^{*}\right\|_{H^{*}}
\end{aligned}
$$

Further, $T$ is injective. For, if $T h_{1}^{*}=T h_{2}^{*}$, then

$$
0=<T h_{1}^{*}-T h_{2}^{*}, v>_{V^{*} \times V}=<h_{1}^{*}-h_{2}^{*}, v>_{H^{*} \times H}
$$

for all $v \in V$ and the statement follows from density of $V$ in $H$. Finally, the image of $T H^{*}$ is dense in $V^{*}$. Indeed, let $v \in V^{* *}$ be such that $<v, T h^{*}>=0$ for all $h^{*} \in H^{*}$. Then, by reflexivity,

$$
0=<v, T h^{*}>_{V^{* *} \times V^{*}}=<T h^{*}, v>_{V^{*} \times V}=<h^{*}, v>_{H^{*} \times H}, \quad h^{*} \in H^{*}
$$

implies $v=0$.
Now, if we identify $H^{*}$ with $H$ by the Riesz theorem and using $T$ as the canonical embedding from $H^{*}$ into $V^{*}$, one writes

$$
V \subset H \simeq H^{*} \subset V^{*}
$$

and the injections are dense and continuous. In such a case we say that $H$ is the pivot space. Note that the scalar product in $H$ coincides with the duality pairing $<\cdot, \cdot>_{V^{*} \times V}$ :

$$
(f, g)_{H}=<f, g>_{V^{*} \times V}, \quad f \in H, g \in V
$$

Remembering now that $V$ is a Hilbert space with scalar product $(\cdot, \cdot)_{V}$ we see that identifying also $V$ with $V^{*}$ would lead to an absurd - we would have $V=H=H^{*}=V^{*}$. Thus, we cannot identify simultaneously both pairs. In such situations it is common to identify the pivot space $H$ with its dual $H^{*}$ bur to leave $V$ and $V^{*}$ as separate spaces with duality pairing being an extension of the scalar product in $H$.

An instructive example is $H=L_{2}([0,1], d x)$ (real) with scalar product

$$
(u, v)=\int_{0}^{1} u(x) v(x) d x
$$

and $V=L_{2}([0,1], w d x)$ with scalar product

$$
(u, v)=\int_{0}^{1} u(x) v(x) w(x) d x
$$

where $w$ is a nonnegative unbounded measurable function. Then it is useful to identify $V^{*}=L_{2}\left([0,1], w^{-1} d x\right)$ and

$$
<f, g>_{V^{*} \times V}=\int_{0}^{1} f(x) g(x) d x \leq \int_{0}^{1} f(x) \sqrt{w(x)} \frac{g(x)}{\sqrt{w(x)}} d x \leq\|f\|_{V}\|g\|_{V^{*}}
$$

### 1.3.2 The Radon-Nikodym theorem

Let $\mu$ and $\nu$ be finite nonnegative measures on the same $\sigma$-algebra in $\Omega$. We say that $\nu$ is absolutely continuous with respect to $\mu$ if every set that has $\mu$-measure 0 also has $\nu$ measure 0 .

Theorem 1.37. If $\nu$ is absolutely continuous with respect to $\mu$ then there is an integrable function $g$ such that

$$
\begin{equation*}
\nu(E)=\int_{E} g d \mu \tag{1.34}
\end{equation*}
$$

for any $\mu$-measurable set $E \subset \Omega$.
Proof. Assume for simplicity that $\mu(\Omega), \nu(\Omega)<\infty$. Let $H=L_{2}(\Omega, d \mu+d \nu)$ on the field of reals. Schwarz inequality shows that if $f \in H$, then $f \in L_{1}(d \mu+$ $d \nu)$, then the linear functional

$$
<x^{*}, f>:=\int_{\Omega} f d \mu
$$

is bounded on $H$. Indeed
$\left|<x^{*}, f>\right| \leq \int_{\Omega} 1 \cdot f d \mu \leq \sqrt{\mu(\Omega)} \sqrt{\int_{\Omega} f^{2} d \mu} \leq \sqrt{\mu(\Omega)} \sqrt{\int_{\Omega} f^{2} d(\mu+\nu)} \leq \sqrt{\mu(\Omega)}\|f\|_{H}$.
Thus, by the Riesz theorem, there is $y \in H$ such that

$$
\int_{\Omega} f d \mu=\int_{\Omega} f y d(\mu+\nu)
$$

Thus we obtain

$$
\int_{\Omega} f(1-y) d \mu=\int_{\Omega} f y d \nu
$$

We claim that $0<y \leq 1$ almost everywhere with respect to $\mu$ (and thus $\nu$ ). Consider the set $F=\{x \in \Omega ; y \leq 0\}$ and $f$ as the characteristic function of $F, f=\chi_{F}$ so that

$$
\int_{F}(1-y) d \mu=\int_{F} y d \nu
$$

If $\mu(F)>0$, then the left hand side is bigger that $\mu(F)>0$ and the right hand side is at most 0 (as it may happen that $\nu(F)=0$ - absolute continuity works only one way). Thus, $\mu(F)=0$ and $y>0 \mu$ (and $\nu$ ) almost everywhere. Let now $E=\{x \in \Omega ; y>0\}$ and $f$ be the characteristic function of $E$ so that

$$
\int_{E}(1-y) d \mu=\int_{E} y d \nu
$$

Now, if $\mu(E)>0$, then the left hand side is strictly negative whereas the right hand side is at least 0 (if $\nu(E)=0$ ). Thus, $\mu(E)=0$ and $y \leq 1 \mu$ (and $\nu$ ) almost everywhere. We can modify $y$ on a $\mu$ measure zero set so that $0<y \leq 1$ everywhere so that

$$
g=\frac{1-y}{y}
$$

is a finite nonnegative function on $\Omega$. Let us denote

$$
E_{n}=\left\{x \in \Omega ; y(x) \geq n^{-1}\right\}
$$

The sequence $\left(E_{n}\right)_{n \in \mathbb{N}}$ is a nested sequence with $\bigcap E_{n}=\emptyset$ as $y$ is positive everywhere. Thus we ca write $\chi_{E_{n}}=y f_{n}$ for some $f_{n} \in H$. Indeed, $0 \leq$ $f_{n} \leq \chi_{E_{n}} / y \leq n$ so that $f$ is bounded and thus square integrable for each $n$. Therefore we can write

$$
\int_{\Omega} \chi_{E_{n}} y^{-1}(1-y) d \mu=\int_{\Omega} \chi_{E_{n}} d \nu
$$

Since $\chi_{E_{n}} \nearrow 1$ everywhere on $\Omega$, using the dominated convergence theorem we obtain that $g=y^{-1}(1-y)$ is integrable on $\Omega$. Taking arbitrary measurable subset $E \subset \Omega$ and its characteristic function, we obtain

$$
\nu(E)=\int_{E} d \nu=\int_{E} g d \mu
$$

### 1.3.3 Projection on a convex set

Corollary 1.38. Let $K$ be a closed convex subset of a real Hilbert space $H$. For any $x \in H$ there is a unique $y \in K$ such that

$$
\begin{equation*}
\|x-y\|=\inf _{z \in K}\|x-z\| \tag{1.35}
\end{equation*}
$$

Moreover, $y \in K$ is a unique solution to the variational inequality

$$
\begin{equation*}
(x-y, z-y) \leq 0 \tag{1.36}
\end{equation*}
$$

for any $z \in K$.
Proof. Let $d=\inf _{z \in K}\|x-z\|$. We can assume $x \notin K$ and so $d>0$. Consider $f(z)=\|x-z\|, z \in K$ and consider a minimizing sequence $\left(z_{n}\right)_{n \in \mathbb{N}}, z_{n} \in K$ such that $d \leq f\left(z_{n}\right) \leq d+1 / n$. By the definition of $f,\left(z_{n}\right)_{n \in \mathbb{N}}$ is bounded and thus it contains a weakly convergent subsequence, say $\left(\zeta_{n}\right)_{n \in \mathbb{N}}$. Since $K$ is closed and convex, by Corollary $1.30, \zeta_{n} \rightharpoonup y \in K$. Further we have
$|(h, x-y)|=\lim _{n \rightarrow \infty}\left|\left(h, x-\zeta_{n}\right)\right| \leq\|h\| \liminf _{n \rightarrow \infty}\left\|x-\zeta_{n}\right\| \leq\|h\| \liminf _{n \rightarrow \infty} d+\frac{1}{n}=\|h\| d$
for any $h \in H$ and thus, taking supremum over $\|h\| \leq 1$, we get $f(y) \leq d$ which gives existence of a minimizer.

To prove equivalence of (1.36) and (1.35) assume first that $y \in K$ satisfies (1.35) and let $z \in K$. Then, from convexity, $v=(1-t) y+t z \in K$ for $t \in[0,1]$ and thus

$$
\|x-y\| \leq\|x-((1-t) y+t z)\|=\|(x-y)-t(z-y)\|
$$

and thus

$$
\|x-y\|^{2} \leq\|x-y\|^{2}-2 t(x-y, z-y)+t^{2}\|z-y\|^{2}
$$

Hence

$$
t\|z-y\|^{2} \geq 2(x-y, z-y)
$$

for any $t \in(0,1]$ and thus, passing with $t \rightarrow 0,(x-y, z-y) \leq 0$. Conversely, assume (1.36) is satisfied and consider

$$
\begin{aligned}
\|x-y\|^{2}-\|x-z\|^{2} & =(x-y, x-y)-(x-z, x-z) \\
& =2(x, z)-2(x, y)+2(y, y)-2(y, z)+2(y, z)-(y, y) \\
& =2(x-y, z-y)-(y-z, y-z) \leq 0
\end{aligned}
$$

hence

$$
\|x-y\| \leq\|x-z\|
$$

for any $z \in K$.

For uniqueness, let $y_{1}, y_{2}$ satisfy

$$
\left(x-y_{1}, z-y_{1}\right) \leq 0, \quad\left(x-y_{2}, z-y_{2}\right) \leq 0, \quad z \in H
$$

Choosing $z=y_{2}$ in the first inequality and $z=y_{1}$ in the second and adding them, we get $\left\|y_{1}-y_{2}\right\|^{2} \leq 0$ which implies $y_{1}=y_{2}$.

We call the operator assigning to any $x \in K$ the element $y \in K$ satisfying (1.35) the projection onto $K$ and denote it by $P_{K}$.

Proposition 1.39. Let $K$ be a nonempty closed and convex set. Then $P_{K}$ is non expansive mapping.

Proof. Let $y_{i}=P_{K} x_{i}, i=1,2$. We have

$$
\left(x_{1}-y_{1}, z-y_{1}\right) \leq 0, \quad\left(x_{2}-y_{2}, z-y_{2}\right) \leq 0, \quad z \in H
$$

so choosing, as before, $z=y_{2}$ in the first and $z=y_{1}$ in the second inequality and adding them together we obtain

$$
\left\|y_{1}-y_{2}\right\|^{2} \leq\left(x_{1}-x_{2}, y_{1}-y_{2}\right)
$$

hence $\left\|P_{K} x_{1}-P_{K} x_{2}\right\| \leq\left\|x_{1}-x_{2}\right\|$.

### 1.3.4 Theorems of Stampacchia and Lax-Milgram

### 1.3.5 Motivation

Consider the Dirichlet problem for the Laplace equation in $\Omega \subset \mathbb{R}^{n}$

$$
\begin{align*}
-\Delta u & =f \quad \text { in } \quad \Omega  \tag{1.37}\\
\left.u\right|_{\partial \Omega} & =0 \tag{1.38}
\end{align*}
$$

Assume that there is a solution $u \in C^{2}(\Omega) \cap C(\bar{\Omega})$. If we multiply (1.37) by a test function $\phi \in C_{0}^{\infty}(\Omega)$ and integrate by parts, then we obtain the problem

$$
\begin{equation*}
\int_{\Omega} \nabla u \cdot \nabla \phi d x=\int_{\Omega} f \phi d x \tag{1.39}
\end{equation*}
$$

Conversely, if $u$ satisfies (1.39), then it is a distributional solution to (1.37).
Moreover, if we consider the minimization problem for

$$
J(u)=\frac{1}{2} \int_{\Omega}|\nabla u|^{2} d x-\int_{\Omega} f u d x
$$

over $K=\left\{u \in C^{2}(\Omega) ;\left.u\right|_{\partial \Omega}=0\right\}$ and if $u$ is a solution to this problem then for any $\epsilon \in \mathbb{R}$ and $\phi \in C_{0}^{\infty}(\Omega)$ we have

$$
J(u+\epsilon \phi) \geq J(u)
$$

then we also obtain (1.39). The question is how to obtain the solution.
In a similar way, we consider the obstacle problem, to minimize $J$ over $K=\left\{u \in C^{2}(\Omega) ;\left.u\right|_{\partial \Omega}=0, u \geq g\right\}$ over some continuous function $g$ satisfying $\left.g\right|_{\partial \Omega}<0$. Note that $K$ is convex. Again, if $u \in K$ is a solution then for any $\epsilon>0$ and $\phi \in K$ we obtain that $u+\epsilon(\phi-u)=(1-\epsilon) u+\epsilon \phi$ is in $K$ and therefore

$$
J(u+\epsilon(\phi-u)) \geq J(u)
$$

Here, we obtain only

$$
\begin{equation*}
\int_{\Omega} \nabla u \cdot \nabla(\phi-u) d x \geq \int_{\Omega} f(\phi-u) d x \tag{1.40}
\end{equation*}
$$

for any $\phi \in K$. For twice differentiable $u$ we obtain

$$
\int_{\Omega} \Delta u(\phi-u) d x \geq \int_{\Omega} f(\phi-u) d x
$$

and choosing $\phi=u+\psi, 0 \leq \psi \in C_{0}^{\infty}(\Omega)$ we get

$$
-\Delta u \geq f
$$

almost everywhere on $\Omega$. As $u$ is continuous, the set $N=\{x \in \Omega ; u(x)>$ $g(x)\}$ is open. Thus, taking $\psi \in C_{0}^{\infty}(N)$, we see that for sufficiently small $\epsilon>0, u \pm \epsilon \phi \in K$. Then, on $N$

$$
-\Delta u=f
$$

Summarizing, for regular solutions the minimizer satisfies

$$
\begin{aligned}
-\Delta u & \geq f \\
u & \geq g \\
(\Delta u+f)(u-g) & =0
\end{aligned}
$$

on $\Omega$.

## Hilbert space theory

We begin with the following definition.
Definition 1.40. Let $H$ be a Hilbert space. A bilinear form $a: H \times H \rightarrow \mathbb{R}$ is said to be
(i) continuous of there is a constant $C$ such that

$$
|a(x, y)| \leq C\|x\|\|y\|, \quad x, y \in H
$$

coercive if there is a constant $\alpha>0$ such that

$$
a(x, x) \geq \alpha\|x\|^{2}
$$

Note that in the complex case, coercivity means $|a(x, x)| \geq \alpha\|x\|^{2}$.
Theorem 1.41. Assume that $a(\cdot, \cdot)$ is a continuous coercive bilinear form on a Hilbert space $H$. Let $K$ be a nonempty closed and convex subset of $H$. Then, given any $\phi \in H^{*}$, there exists a unique element $x \in K$ such that for any $y \in K$

$$
\begin{equation*}
a(x, y-x) \geq<\phi, y-x>_{H^{*} \times H} \tag{1.41}
\end{equation*}
$$

Moreover, if $a$ is symmetric, then $x$ is characterized by the property

$$
\begin{equation*}
x \in K \quad \text { and } \quad \frac{1}{2} a(x, x)-<\phi, x>_{H^{*} \times H}=\min _{y \in K} \frac{1}{2} a(y, y)-<\phi, y>_{H^{*} \times H} . \tag{1.42}
\end{equation*}
$$

Proof. First we note that from Riesz theorem, there is $f \in H$ such that $<\phi, y>_{H^{*} \times H}=(f, y)$ for all $y \in H$. Now, if we fix $x \in H$, then $y \rightarrow a(x, y)$ is a continuous linear functional on $H$. Thus, again by the Riesz theorem, there is an operator $A: H \rightarrow H$ satisfying $a(x, y)=(A x, y)$. Clearly, $A$ is linear and satisfies

$$
\begin{align*}
\|A x\| & \leq C\|x\|  \tag{1.43}\\
(A x, x) & \geq \alpha\|x\|^{2} \tag{1.44}
\end{align*}
$$

Indeed,

$$
\|A x\|=\sup _{\|y\|=1}|(A x, y)| \leq C\|x\| \sup _{\|y\|=1}\|y\|
$$

and (1.44) is obvious.
Problem (1.41) amounts to finding $x \in K$ satisfying, for all $y \in K$,

$$
\begin{equation*}
(A x, y-x) \geq(f, y-x) \tag{1.45}
\end{equation*}
$$

Let us fix a constant $\rho$ to be determined later. Then, multiplying both sides of (1.45) by $\rho$ and moving to one side, we find that (1.45) is equivalent to

$$
\begin{equation*}
(\rho f-\rho A x+x-x, y-x) \leq 0 \tag{1.46}
\end{equation*}
$$

Here we recognize the equivalent formulation of the projection problem (1.36), that is, we can write

$$
\begin{equation*}
x=P_{K}(\rho f-\rho A x+x) \tag{1.47}
\end{equation*}
$$

This is a fixed point problem for $x$ in $K$. Denote $S y=P_{K}(\rho f-\rho A y+y)$ Clearly $S: K \rightarrow K$ as it is a projection onto $K$ and $K$, being closed, is a complete metric space in the metric induced from $H$. Since $P_{K}$ is nonexpansive, we have

$$
\left\|S y_{1}-S y_{2}\right\| \leq\left\|\left(y_{1}-y_{2}\right)-\rho\left(A y_{1}-A y_{2}\right)\right\|
$$

and thus

$$
\begin{aligned}
\left\|S y_{1}-S y_{2}\right\|^{2} & =\left\|y_{1}-y_{2}\right\|^{2}-2 \rho\left(A y_{1}-A y_{2}, y_{1}-y_{2}\right)+\rho^{2}\left\|A y_{1}-A y_{2}\right\|^{2} \\
& \leq\left\|y_{1}-y_{2}\right\|^{2}\left(1-2 \rho \alpha+\rho^{2} C^{2}\right)
\end{aligned}
$$

We can choose $\rho$ in such a way that $k^{2}=1-2 \rho \alpha+\rho^{2} C^{2}<1$ we see that $S$ has a unique fixed point in $K$.

Assume now that $a$ is symmetric. Then $(x, y)_{1}=a(x, y)$ defines a new scalar product which defines an equivalent norm $\|x\|_{1}=\sqrt{a(x, x)}$ on $H$. Indeed, by continuity and coerciveness

$$
\|x\|_{1}=\sqrt{a(x, x)} \leq \sqrt{C}\|x\|
$$

and

$$
\|x\|=\sqrt{a(x, x)} \geq \sqrt{\alpha}\|x\|
$$

Using again Riesz theorem, we find $g \in H$ such that

$$
<\phi, y>_{H^{*} \times H}=a(g, y)
$$

and then (1.41) amounts to finding $x \in K$ such that

$$
a(g-x, y-x) \leq 0
$$

for all $y \in K$ but this is nothing else but finding projection $x$ onto $K$ with respect to the new scalar product. Thus, there is a unique $x \in K$

$$
\sqrt{a(g-x, g-x)}=\min _{y \in K} \sqrt{a(g-x, g-x)} .
$$

However, expanding, this is the same as finding minimum of the function
$y \rightarrow a(g-y, g-y)=a(g, g)+a(y, y)-2 a(g, y)=a(y, y)-2<\phi, y>_{H^{*} \times H}+a(g, g)$.
Taking into account that $a(g, g)$ is a constant, we see that $x$ is the unique minimizer of

$$
y \rightarrow \frac{1}{2} a(y, y)-<\phi, y>_{H^{*} \times H}
$$

Corollary 1.42. Assume that $a(\cdot, \cdot)$ is a continuous coercive bilinear form on a Hilbert space $H$. Then, given any $\phi \in H^{*}$, there exists a unique element $x \in H$ such that for any $y \in H$

$$
\begin{equation*}
a(x, y)=<\phi, y>_{H^{*} \times H} \tag{1.48}
\end{equation*}
$$

Moreover, if $a$ is symmetric, then $x$ is characterized by the property

$$
\begin{equation*}
x \in H \quad \text { and } \quad \frac{1}{2} a(x, x)-<\phi, x>_{H^{*} \times H}=\min _{y \in H} \frac{1}{2} a(y, y)-<\phi, y>_{H^{*} \times H} \tag{1.49}
\end{equation*}
$$

Proof. We use the Stampacchia theorem with $K=H$. Then there is a unique element $x \in H$ satisfying

$$
a(x, y-x) \geq<\phi, y-x>_{H^{*} \times H}
$$

Using linearity, this must hold also for

$$
a(x, t y-x) \geq<\phi, t y-x>_{H^{*} \times H}
$$

for any $t \in R, v \in H$. Factoring out $t$, we find

$$
t a\left(x, y-x t^{-1}\right) \geq t<\phi, y-x t^{-1}>_{H^{*} \times H}
$$

and passing with $t \rightarrow \pm \infty$, we obtain

$$
a(x, y) \geq<\phi, y>_{H^{*} \times H}, \quad a(x, y) \leq<\phi, y>_{H^{*} \times H}
$$

Remark 1.43. Elementary proof of the Lax-Milgram theorem. As we noted earlier

$$
a(x, y)=<\phi, y>_{H^{*} \times H}
$$

can be written as the equation

$$
(A x, y)=(f, y)
$$

for any $y \in H$, where $A: H \rightarrow H,\|A x\| \leq C\|x\|$ and $(A x, x) \geq \alpha\|x\|^{2}$. From the latter, $A x=0$ implies $x=0$, hence $A$ is injective. Further, if $y=A x$, $x=A^{-1} y$ and

$$
\|x\|^{2}=\left\|A^{-1} y\right\|\|x\| \leq \alpha^{-1}(y, x) \leq \alpha^{-1}\|y\|\|x\|
$$

so $A^{-1}$ is bounded. This shows that the range of $A, R(A)$, is closed. Indeed, if $\left(y_{n}\right)_{n \in \mathbb{N}}, y_{n} \in R(A), y_{n} \rightarrow y$, then $\left(y_{n}\right)_{n \in \mathbb{N}}$ is Cauchy, but then $\left(x_{n}\right)_{n \in \mathbb{N}}$, $x_{n}=A^{-1}$ is also Cauchy and thus converges to some $x \in A$. But then, from continuity of $A, A x=y$. On the other hand, $R(A)$ is dense. For, if for some $v \in H$ we have $0=(A x, v)$ for any $x \in H$, we can take $v=x$ and

$$
0=(A v, v) \geq \alpha\|v\|^{2}
$$

so $v=0$ and so $R(A)$ is dense.

### 1.3.6 Dirchlet problem

Let us recall the variational formulation of the Dirichlet problem: find $u \in$ ? such that

$$
\begin{equation*}
\int_{\Omega} \nabla u \cdot \nabla \phi d x=\int_{\Omega} f \phi d x \tag{1.50}
\end{equation*}
$$

for all $C_{0}^{\infty}(\Omega)$. We also recall the associated minimization problem for

$$
\begin{equation*}
J(u)=\frac{1}{2} \int_{\Omega}|\nabla u|^{2} d x-\int_{\Omega} f u d x \tag{1.51}
\end{equation*}
$$

over some closed subspace $K=\{u \in ?\}$.
Let us consider the space $H=L_{2}(\Omega), \Omega \subset \mathbb{R}^{n}$ bounded, with the scalar product

$$
(u, v)_{0}=\int_{\Omega} u(x) v(x) d x
$$

We know that ${\overline{C_{0}^{\infty}(\Omega)}}^{H}=H$. The relation (1.50) suggests that we should consider another scalar product, initially on $C_{0}^{\infty}(\Omega)$, given by

$$
(u, v)_{0,1}=\int_{\Omega} \nabla u(x) \nabla v(x) d x
$$

Note that due to the fact that $u, v$ have compact supports, this is a well defined scalar product as

$$
0=(u, u)_{0,1}=\int_{\Omega}|\nabla u(x)|^{2} d x
$$

implies $u_{x_{i}}=0$ for all $x_{i}, i=1, \ldots, n$ hence $u=$ const and thus $u \equiv 0$. Note that this is not a scalar product on a space $C^{\infty}(\bar{\Omega})$.

A fundamental role in the theory is played by the Zaremba - PoincarèFriedrichs lemma.
Lemma 1.44. There is a constant $d$ such that for any $u \in C_{0}^{\infty}(\Omega)$

$$
\begin{equation*}
\|u\|_{0} \leq d\|u\|_{0,1} \tag{1.52}
\end{equation*}
$$

Proof. Let $R$ be a box $\left[a_{1}, b_{1}\right] \times \ldots \times\left[a_{n}, b_{n}\right]$ such that $\bar{\Omega} \subset R$ and extend $u$ by zero to $R$. Since $u$ vanishes at the boundary of $R$, for any $\mathbf{x}=\left(x_{1}, \ldots, x_{n}\right)$ we have

$$
u(\mathbf{x})=\int_{a_{i}}^{x_{i}} u_{x_{i}}\left(x_{1}, \ldots, t, \ldots\right) d t
$$

and, by Schwarz inequality,

$$
\begin{aligned}
u^{2}(\mathbf{x}) & =\left(\int_{a_{i}}^{x_{i}} u_{x_{i}}\left(x_{1}, \ldots, t, \ldots, x_{n}\right) d t\right)^{2} \leq\left(\int_{a_{i}}^{x_{i}} 1 d t\right)\left(\int_{a_{i}}^{x_{i}} u_{x_{i}}^{2}\left(x_{1}, \ldots, t, \ldots, x_{n}\right) d t\right) \\
& \leq\left(b_{i}-a_{i}\right) \int_{a_{i}}^{b_{i}} u_{x_{i}}^{2}\left(x_{1}, \ldots, t, \ldots, x_{n}\right) d t
\end{aligned}
$$

for any $\mathbf{x} \in R$. Integrating over $R$ we obtain

$$
\int_{R} u^{2}(\mathbf{x}) d \mathbf{x} \leq\left(b_{i}-a_{i}\right)^{2} \int_{R} u_{x_{i}}^{2}(\mathbf{x}) d \mathbf{x} .
$$

This can be re-written

$$
\int_{\Omega} u^{2}(\mathbf{x}) d \mathbf{x} \leq\left(b_{i}-a_{i}\right)^{2} \int_{\Omega} u_{x_{i}}^{2}(\mathbf{x}) d \mathbf{x} \leq c \int_{\Omega}|\nabla u(\mathbf{x})|^{2} d \mathbf{x}
$$

We see that the lemma remains valid if $\Omega$ is bounded just in one direction. Let us define ${ }_{W}^{\mathrm{o}}{ }_{2}^{1}(\Omega)$ as the completion of $C_{0}^{\infty}(\Omega)$ in the norm $\|\cdot\|_{0,1}$. We have
Theorem 1.45. The space ${ }^{\circ}{ }_{2}^{1}(\Omega)$ is a separable Hilbert space which can be identified with a subspace continuously and densely embedded in $L_{2}(\Omega)$. Every $v \in \stackrel{\circ}{W}_{2}^{1}(\Omega)$ has generalized derivatives $D_{x_{i}} v \in L_{2}(\Omega)$. Furthermore, the distributional integration by parts formula

$$
\begin{equation*}
\int_{\Omega} D_{x_{i}} v u d \mathbf{x}=-\int_{\Omega} v D_{x_{i}} u d \mathbf{x} \tag{1.53}
\end{equation*}
$$

is valid for any $u, v \in \stackrel{\circ}{W}_{2}^{1}(\Omega)$.
Proof. . The completion in the scalar product gives a Hilbert space. By Lemma 1.44 , every equivalence class of the completion in the norm $\|\cdot\|_{0,1}$ is also an equivalence class in $\|\cdot\|_{0}$ and thus can be identified with the element of $\overline{C_{0}^{\infty}(\Omega)}\left\|^{\|}\right\|_{0}$ and thus with an element $v \in L_{2}(\Omega)$. This identification is one-to-one. Density follows from $C_{0}^{\infty}(\Omega) \subset{ }_{W}^{\circ} \frac{1}{2}(\Omega) \subset L_{2}(\Omega)$ and continuity of injection from Lemma 1.44.

If $\left(v_{n}\right)_{n \in \mathbb{N}}$ of $C_{0}^{\infty}(\Omega)$ functions converges to $v \in W_{2}^{1}(\Omega)$ in $\|\cdot\|_{0,1}$, then $v_{n} \rightarrow v$ in $L_{2}(\Omega)$ and $D_{x_{i}} v_{n} \rightarrow v^{i}$ in $L_{2}(\Omega)$ for some functions $v^{i} \in L_{2}(\Omega)$. Taking arbitrary $\phi \in C_{0}^{\infty}(\Omega)$, we obtain

$$
\int_{\Omega} D_{x_{i}} v_{n} \phi d \mathbf{x}=-\int_{\Omega} v_{n} D_{x_{i}} \phi d \mathbf{x}
$$

and we can pass to the limit

$$
\int_{\Omega} v^{i} \phi d \mathbf{x}=-\int_{\Omega} v D_{x_{i}} \phi d \mathbf{x}
$$

showing that $v^{i}=D_{x_{i}} v$ in generalized sense. Furthermore, we can pass to the limit in $\|\cdot\|_{0,1}$ with $\phi \rightarrow u \in W_{2}^{\mathrm{o}}{ }_{2}^{1}(\Omega)$ and, by the above, $D_{x_{i}} \phi \rightarrow D_{x_{i}} u$ in
$L_{2}(\Omega)$, giving (1.53). This also shows that ${ }^{\circ}{ }_{2}^{1}(\Omega)$ can be identified with a closed subspace of $\left(L_{2}(\Omega)\right)^{n}$ (the graph of gradient) and thus it is a separable space.

Consider now on ${ }_{W}^{\mathrm{o}}{ }_{2}^{1}(\Omega)$ the bilinear form

$$
a(u, v)=\int_{\Omega} \nabla u \nabla v d \mathbf{x}
$$

Clearly, by Schwarz inequality

$$
|a(u, v)| \leq\|u\|_{0,1}\|v\|_{0,1}
$$

and

$$
a(u, u)=\int_{\Omega} \nabla u \nabla u d \mathbf{x}=\|u\|_{0,1}^{2}
$$

and thus $a$ is a continuous and coercive bilinear form on ${ }_{W}^{\circ}{ }_{2}^{1}(\Omega)$. Thus, if we take $f \in\left(\stackrel{\circ}{W}_{2}^{1}(\Omega)\right)^{*} \supset L_{2}(\Omega)$ then there is a unique $u \in \stackrel{\circ}{W}_{2}^{1}(\Omega)$ satisfying

$$
\int_{\Omega} \nabla u \nabla v d \mathbf{x}=<f, v>_{\left(W_{2}^{\mathrm{o}}(\Omega)\right)^{*} \times W_{2}^{\mathrm{o}}(\Omega)}
$$

for any $v \in W_{2}^{1}(\Omega)$ or, equivalently, minimizing the functional

$$
J(v)=\frac{1}{2} \int_{\Omega}|\nabla v|^{2} d \mathbf{x}-<f, v>_{\left(W_{2}^{\mathfrak{A}}(\Omega)\right)^{*} \times W_{2}^{\mathcal{A}}(\Omega)}
$$

over $K=\stackrel{o}{W}_{2}^{1}(\Omega)$.
The question is what this solution represents. Clearly, taking $v \in C_{0}^{\infty}(\Omega)$ we obtain

$$
-\Delta u=f
$$

in the sense of distribution. However, to get a deeper understanding of the meaning of the solution, we have investigate the structure of ${ }_{W}^{\circ}{ }_{2}^{1}(\Omega)$.

### 1.3.7 Sobolev spaces

Let $\Omega$ be a nonempty open subset of $\mathbb{R}^{n}, n \geq 1$ and let $m \in \mathbb{N}$. The Sobolev space $W_{2}^{m}(\Omega)$ consists of all $u \in L_{2}(\Omega)$ for which all generalized derivatives $D^{\alpha} u$ exist and belong to $L_{2} . W_{2}^{m}(\Omega)$ is equipped with the scalar product

$$
\begin{equation*}
(u, v)_{m}=\sum_{|\alpha| \leq m} \int_{\Omega} D^{\alpha} u D^{\alpha} v d \mathbf{x} \tag{1.54}
\end{equation*}
$$

In particular,

$$
(u, v)_{1}=\int_{\Omega} u v+\nabla u \nabla v d \mathbf{x}
$$

We obtain
Proposition 1.46. The space $W_{2}^{m}(\Omega)$ is a separable Hilbert space.
Proof. The proof follows since the generalized differentiation is a closed operator in $L_{2}(\Omega)$.

We note that ${ }_{W}^{\circ}{ }_{2}^{1}(\Omega)$ is a closed subspace of $W_{2}^{1}(\Omega)$ as the norms $\|\cdot\|_{0,1}$ and $\|\cdot\|_{1}$ coincide there.

We shall focus on the case $m=1$. A workhorse of the theory is the Friedrichs lemma.
Lemma 1.47. Let $u \in W_{2}^{1}(\Omega)$. Then there exists a sequence $\left(u_{k}\right)_{k \in \mathbb{N}}$ from $C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ such that

$$
\begin{equation*}
\left.u_{k}\right|_{\Omega} \rightarrow u \quad \text { in } \quad L_{2}(\Omega) \tag{1.55}
\end{equation*}
$$

and for any $\Omega^{\prime} \Subset \Omega$

$$
\begin{equation*}
\left.\nabla u_{k}\right|_{\Omega^{\prime}} \rightarrow \nabla u \quad \text { in } \quad L_{2}\left(\Omega^{\prime}\right) \tag{1.56}
\end{equation*}
$$

If $\Omega=\mathbb{R}^{n}$, then both convergences occur in $\mathbb{R}^{n}$.
Proof. Set

$$
u^{e}(x)= \begin{cases}u(x) & \text { for } x \in \Omega \\ 0 & \text { for } x \notin \Omega\end{cases}
$$

and define $v_{\epsilon}=u^{e} * \omega_{\epsilon}$. We know $v_{\epsilon} \in C^{\infty}\left(\mathbb{R}^{n}\right)$ and $v_{\epsilon} \rightarrow u$ in $L_{2}(\Omega)$. Let us take $\Omega^{\prime} \Subset \Omega$ and fix a function $\alpha \in C_{0}^{\infty}(\Omega)$ which equals 1 on a neighbourhood of $\Omega^{\prime}$. Then, for sufficiently small $\epsilon$, we have

$$
\omega_{\epsilon} *(\alpha u)^{\epsilon}=\omega_{\epsilon} * u^{\epsilon}
$$

on $\Omega^{\prime}$. Then, by Proposition 1.8,

$$
\partial_{x_{j}}\left(\omega_{\epsilon} *(\alpha u)^{\epsilon}\right)=\omega_{\epsilon} *\left(\alpha \partial_{x_{j}} u+\partial_{x_{j}} \alpha u\right)^{e}
$$

hence

$$
\partial_{x_{j}}\left(\omega_{\epsilon} *(\alpha u)^{\epsilon}\right) \rightarrow\left(\alpha \partial_{x_{j}} u+\partial_{x_{j}} \alpha u\right)^{e}
$$

in $L_{2}(\Omega)$ and, in particular,

$$
\partial_{x_{j}}\left(\omega_{\epsilon} *(\alpha u)^{\epsilon}\right) \rightarrow p_{j} u
$$

in $L_{2}\left(\Omega^{\prime}\right)$. But on $\Omega^{\prime}$ we can discard $\alpha$ to get

$$
\partial_{x_{j}}\left(\omega_{\epsilon} * u^{\epsilon}\right) \rightarrow p_{j} u
$$

If $v_{k}$ do not have compact support (e.g. when $\Omega$ is not bounded), then we multiply $v_{k}$ by a sequence of smooth cut-off functions $\zeta_{k}=\zeta(x / k)$ where $\zeta(x)=1$ for $|x| \leq 1$ and $\zeta(x)=0$ for $|x| \geq 2$.

As an immediate application we show
Proposition 1.48. (i) Let $u, v \in W_{2}^{1}(\Omega) \cap L_{\infty}(\Omega)$. Then $u v \in W_{2}^{1}(\Omega) \cap$ $L_{\infty}(\Omega)$ with

$$
\begin{equation*}
\partial_{x_{j}}(u v)=\partial_{j} u v+u \partial_{x_{j}} v, \quad i=1, \ldots, n \tag{1.57}
\end{equation*}
$$

(ii) Let $\Omega, \Omega_{1}$ be two open sets in $\mathbb{R}^{n}$ and let $H: \Omega_{1} \rightarrow \Omega$ be a $C^{1}(\bar{\Omega})$ diffeomorphism. If $u \in W_{2}^{1}(\Omega)$ then $u \circ H \in W_{2}^{1}\left(\Omega^{\prime}\right)$ and

$$
\begin{equation*}
\left.\int_{\Omega_{1}}(u \circ H) \partial_{y_{j}} \phi d \mathbf{y}=-\int_{\Omega_{1}} \sum_{i=1}^{n}\left(\partial_{x_{i}} u \circ H\right)\right) \partial_{y_{j}} H_{i} \phi d \mathbf{y} \tag{1.58}
\end{equation*}
$$

Proof. Using Friedrichs lemma, we find sequences $\left(u_{k}\right)_{k \in \mathbb{N}},\left(v_{k}\right)_{k \in \mathbb{N}}$ in $C_{0}^{\infty}(\Omega)$ such that

$$
u_{k} \rightarrow u, \quad v_{k} \rightarrow v
$$

in $L_{2}(\Omega)$ and for any $\Omega^{\prime} \Subset \Omega$ we have

$$
\nabla u_{k} \rightarrow \nabla u, \quad \nabla v_{k} \rightarrow \nabla v
$$

in $L_{2}\left(\Omega^{\prime}\right)$. Moreover, from the construction of the mollifiers we get

$$
\left\|u_{k}\right\|_{L_{\infty}(\Omega)} \leq\|u\|_{L_{\infty}(\Omega)} \quad\left\|v_{k}\right\|_{L_{\infty}(\Omega)} \leq\|v\|_{L_{\infty}(\Omega)}
$$

On the other hand

$$
\int_{\Omega} u_{k} v_{k} \partial_{x_{j}} \phi d \mathbf{x}=-\int_{\Omega}\left(\partial_{j} u_{k} v_{k}+u_{k} \partial_{j} v_{k}\right) \phi d \mathbf{x}
$$

for any $\phi \in C_{0}^{\infty}(\Omega)$. Thanks to the compact support of $\phi$, the integration actually occurs over compact subsets of $\Omega$ and we can use $L_{2}$ convergence of $\nabla u_{k}, \nabla v_{k}$. Thus

$$
\int_{\Omega} u v \partial_{x_{j}} \phi d \mathbf{x}=-\int_{\Omega}\left(\partial_{x_{j}} u v+u \partial_{x_{j}} v\right) \phi d \mathbf{x}
$$

and the fact that $u v \in W_{2}^{1}(\Omega)$ follows from $\partial_{x_{j}} u, \partial_{x_{j}} v \in W_{2}^{1}(\Omega)$ and $u, v \in L_{\infty}(\Omega)$. The proof of the second statement follows similarly. We select sequence $\left(u_{k}\right)_{k \in \mathbb{N}}$ as above; then clearly $u_{k} \circ H \rightarrow u \circ H$ in $L_{2}\left(\Omega_{1}\right)$ and

$$
\left(\partial_{x_{i}} u_{k} \circ H\right) \partial_{y_{j}} H_{i} \rightarrow\left(\partial_{x_{i}} u \circ H\right) \partial_{y_{j}} H_{i}
$$

in $L_{2}\left(\Omega_{1}^{\prime}\right)$ for any $\Omega_{1}^{\prime} \Subset \Omega$. For any $\psi \in C_{0}^{\infty}\left(\Omega_{1}\right)$ we get

$$
\int_{\Omega_{1}}\left(u_{k} \circ H\right) \partial_{y_{j}} \phi d \mathbf{y}=-\int_{\Omega_{1}} \sum_{i=1}^{k}\left(\partial_{x_{i}} u_{k} \circ H\right) \partial_{y_{j}} H_{i} \phi d \mathbf{y}
$$

and in the limit we obtain (1.58).

Sometimes it will be necessary to indicate the domain of the definition of a Sobolev space. Then we use the $\|\cdot\|_{0, \Omega}$ to denote the norm in $L_{2}(\Omega)$ and analogous convention is used for the Sobolev space norms.
Proposition 1.49. The following properties are equivalent:
(i) $u \in W_{2}^{1}(\Omega)$,
(ii) there is $C$ such that for any $\phi \in C_{0}^{\infty}(\Omega)$ and $i=1, \ldots, n$

$$
\begin{equation*}
\left|\int_{\Omega} u \partial_{i} \phi d \mathbf{x}\right| \leq C\|\phi\|_{0} \tag{1.59}
\end{equation*}
$$

(iii) there is a constant $C$ such that for any $\Omega^{\prime} \Subset \Omega$ and all $\mathbf{h} \in \mathbb{R}^{n}$ with $|\mathbf{h}| \leq \operatorname{dist}\left(\Omega^{\prime}, \partial \Omega\right)$ we have

$$
\begin{equation*}
\left\|\tau_{h} u-u\right\|_{0, \Omega^{\prime}} \leq C|\mathbf{h}| \tag{1.60}
\end{equation*}
$$

where $\left(\tau_{h} u\right)(\mathbf{x})=u(\mathbf{x}+\mathbf{h})$. In particular, if $\Omega=\mathbb{R}^{n}$, then

$$
\begin{equation*}
\left\|\tau_{h} u-u\right\|_{0} \leq|\mathbf{h}|\|\nabla u\|_{0} \tag{1.61}
\end{equation*}
$$

Proof. $(i) \Rightarrow(i i)$ follows from the definition.
(ii) $\Rightarrow(i)$. Eqn. (1.59) shows that

$$
\phi \rightarrow \int_{\Omega} u \partial_{i} \phi d \mathbf{x}
$$

extends to a bounded functional on $L_{2}(\Omega)$ and thus there is $v_{i} \in L_{2}(\Omega)$ such that

$$
\int_{\Omega} u \partial_{i} \phi d \mathbf{x}=-\int_{\Omega} v_{i} \phi d \mathbf{x}
$$

for any $\phi \in C_{0}^{\infty}(\Omega)$.
(i) $\Rightarrow$ (iii). Let us take $u \in C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$. For $\mathbf{x}, \mathbf{h} \in \mathbb{R}^{n}$ and $t \in \mathbb{R}$ we define

$$
v(t)=u(\mathbf{x}+t \mathbf{h})
$$

Then $v^{\prime}(t)=\mathbf{h} \nabla u(\mathbf{x}+t \mathbf{h})$ and

$$
u(\mathbf{x}+\mathbf{h})-u(\mathbf{x})=v(1)-v(0)=\int_{0}^{1} \mathbf{h} \nabla u(\mathbf{x}+t \mathbf{h}) d t
$$

Hence

$$
\left|\tau_{h} u(\mathbf{x})-u(\mathbf{x})\right|^{2} \leq|\mathbf{h}|^{2} \int_{0}^{1}|\nabla u(\mathbf{x}+t \mathbf{h})|^{2} d t
$$

so that

$$
\begin{aligned}
\int_{\Omega^{\prime}}\left|\tau_{h} u(\mathbf{x})-u(\mathbf{x})\right|^{2} d \mathbf{x} & \leq|\mathbf{h}|^{2} \int_{0}^{1}\left(\int_{\Omega^{\prime}}|\nabla u(x+t h)|^{2} d \mathbf{x}\right) d t \\
& =|\mathbf{h}|^{2} \int_{0}^{1}\left(\int_{\Omega^{\prime}+t \mathbf{h}}|\nabla u(\mathbf{y})|^{2} d \mathbf{x}\right) d t .
\end{aligned}
$$

If $|\mathbf{h}|<\operatorname{dist}\left(\Omega^{\prime}, \partial \Omega\right)$, then there is $\Omega^{\prime \prime}$ such that $\Omega^{\prime}+t \mathbf{h} \subset \Omega^{\prime \prime} \Subset \Omega$ for all $t \in[0,1]$ and thus

$$
\int_{\Omega^{\prime}}\left|\tau_{h} u(\mathbf{x})-u(\mathbf{x})\right|^{2} d \mathbf{x} \leq|\mathbf{h}|^{2} \int_{\Omega^{\prime \prime}}|\nabla u(\mathbf{y})|^{2} d \mathbf{x}
$$

which gives (1.60) for $u \in C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$. Let $u \in W_{2}^{1}(\Omega)$. Then, by the Friedrichs lemma, we find $\left(u_{k}\right)_{k \in \mathbb{N}}, u_{k} \in C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ such that $u_{k} \rightarrow u$ in $L_{2}(\Omega)$ and $\nabla u_{k} \rightarrow \nabla u$ in $L_{2}\left(\Omega^{\prime}\right)$ for any $\Omega^{\prime} \Subset \Omega$. Noting that $\tau_{h} u_{k} \rightarrow \tau_{h} u$ in $L_{2}\left(\Omega^{\prime}\right)$ we can pass to the limit above, obtaining,

$$
\left\|\tau_{h} u-u\right\|_{0, \Omega^{\prime}} \leq|\mathbf{h}| \sqrt{\int_{\Omega^{\prime \prime}}|\nabla u(\mathbf{y})|^{2} d \mathbf{x}} \leq C|\mathbf{h}|
$$

If $\Omega=\mathbb{R}^{n}$, then in all calculations above we can replace $\Omega^{\prime}, \Omega^{\prime \prime}$ by $\mathbb{R}^{n}$.
(iii) $\Rightarrow$ (ii). If (1.60) holds then, taking $\Omega^{\prime} \Subset \Omega, \phi \in C_{0}^{\infty}(\Omega)$ with $\operatorname{supp} \phi \subset$ $\Omega^{\prime}$ and $|\mathbf{h}|<\operatorname{dist}\left(\Omega^{\prime}, \partial \Omega\right)$, we obtain

$$
\left|\int_{\Omega}\left(\tau_{h} u-u\right) \phi d \mathbf{x}\right| \leq C|\mathbf{h}|\|\phi\|_{0}
$$

On the other hand

$$
\int_{\Omega}\left(\tau_{h} u-u\right)(\mathbf{x}) \phi(\mathbf{x}) d \mathbf{x}=\int_{\Omega} u(\mathbf{y})\left(\tau_{-h} \phi-\phi\right)(\mathbf{y}) d \mathbf{y}
$$

so

$$
\int_{\Omega} u \frac{\left(\tau_{-h} \phi-\phi\right)}{|\mathbf{h}|} d \mathbf{y} \leq C\|\phi\|_{0}
$$

Choosing $\mathbf{h}=t \mathbf{e}_{i}, i=1, \ldots, n$ and passing to the limit with $t \rightarrow 0$, we obtain (1.59).

### 1.3.8 Localization and flattening of the boundary

Assume that $\Omega$ is an open, bounded set with boundary $\partial \Omega$ which is an $n-1$ dimensional $C^{m}$ manifold; further assume that that $\Omega$ lies locally at one side of the boundary. Denote $Q=\left\{\mathbf{y} \in \mathbb{R}^{n} ;\left|y_{i}\right|<1, i=1, \ldots, n\right\}, Q_{0}=\{\mathbf{y} \in$ $\left.Q ; y_{n}=0\right\}$ and $Q_{+}=\left\{\mathbf{y} \in Q ; x_{n}>0\right\}$. Then we have a finite local atlas on $\partial \Omega$, that is, a finite collection $\left\{B_{j}, H^{j} j\right\}_{1 \leq j \leq N}$ where $B_{j}$ are open sets covering $\partial \Omega, H^{j}: Q \rightarrow B_{j}$ are $C^{m}$ diffeomorphisms with positive Jacobians which are bijections of $Q, Q_{0}$ and $Q_{+}$onto $B_{j}, B_{j} \cap \partial \Omega$ and $B_{j} \cap \Omega$, respectively.

Given the local atlas $\left\{B_{j}, H^{j}\right\}_{1 \leq j \leq N}$, we construct a finite open subcover $\left\{G_{j}\right\}_{1 \leq j \leq N}$ in such a way that $G_{j} \Subset B_{j}$ and $\partial \Omega \subset \bigcup_{j=1}^{N} G_{j}$. In fact, we can take $G_{j}=B_{j}^{k}$ where $B_{j}^{k}=\left\{\mathbf{x} \in B_{j} ; \operatorname{dist}(\mathbf{x}, \partial \Omega>1 / k\}\right.$ for some $k$. Indeed, suppose it is impossible, then for any $k$ there is $\mathbf{x}_{k} \in \partial \Omega$ such that $\mathbf{x}_{k} \notin \bigcup_{j=1}^{N} B_{j}^{k}$. From compactness of $\partial \Omega$ we obtain an accumulation point $\mathbf{x} \in \partial \Omega$. Hence $\mathbf{x} \in B_{j}$ for some $j$ and thus $x \in B_{j}^{k}$ for sufficiently large $k$. This contradicts the construction that $x$ is an accumulation point of points which are outside $\bigcup_{j=1}^{N} B_{j}^{k}$. Defining $G_{0}=\Omega \backslash \bigcup_{j=1}^{N} \bar{G}_{j}$ we further get an open set $G_{0}$ with $\bar{G}_{0} \subset \Omega$. Thus

$$
\bar{\Omega} \subset \Omega \cup \bigcup_{j=1}^{N} G_{j}, \quad \Omega \subset \bigcup_{j=0}^{N} \bar{G}_{j}
$$

Now, we choose $\alpha_{j} \in C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ satisfying $0 \leq \alpha \leq 1, \operatorname{supp} \alpha_{j} \subset B_{j}$ and $\alpha_{j}=1$ on $\bar{G}_{j}$. Further, $\alpha \in C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ satisfies

$$
\operatorname{supp} \alpha \subset \Omega \cup \bigcup_{j=1}^{N} G_{j}, \quad 0 \leq \alpha \leq 1, \quad \alpha=1 \text { on } \bar{\Omega} .
$$

Then define

$$
\beta_{j}(\mathbf{x})=\frac{\alpha(\mathbf{x}) \alpha_{j}(\mathbf{x})}{\sum_{k=0}^{N} \alpha_{k}(\mathbf{x})}
$$

for $\mathbf{x} \in \bigcup_{j=0}^{N} \bar{G}_{j}$ and $\beta_{j}(\mathbf{x})=0$ for $\mathbf{x} \in \mathbb{R}^{n} \backslash \bigcup_{j=0}^{N} \bar{G}_{j}$. We note that each $\beta_{j}$ is well defined. Indeed, at least one $\alpha_{j}(\mathbf{x})$ is equal 1 on $\bigcup_{j=0}^{N} \bar{G}_{j}$ so that the denominator is at least 1 there. On the other hand, $\alpha$ vanishes outside a compact set contained in $\bigcup_{j=0}^{N} G_{j}$. Hence, $\beta_{j} \in C_{0}^{\infty}\left(\mathbb{R}^{n}\right), \operatorname{supp} \beta_{j} \subset B_{j}$, $\beta_{j} \geq 0$ and

$$
\sum_{j=0}^{N} \beta_{j}(\mathbf{x})=1
$$

for $\mathbf{x} \in \bar{\Omega}$.
We call the collection $\left\{\beta_{j}\right\}_{j=0}^{N}$ a partition of unity subordinated to the open cover $\left\{G_{j}\right\}_{j=0}^{N}$ of $\Omega$ and $\left\{\beta_{j}\right\}_{j=1}^{N}$ a partition of unity subordinated to the open cover $\left\{G_{j}\right\}_{j=1}^{N}$ of $\Omega$ of $\partial \Omega$.

Suppose now we have $u \in W_{2}^{1}(\Omega)$. Then $u=\sum_{j=0}^{N} \beta_{j} u$ on $\Omega$ and, by Proposition 1.48 (i), $\beta_{j} u \in W_{2}^{1}\left(\Omega \cap G_{j}\right), j=1, \ldots, N$. Using Proposition 1.48 (ii) we see that for each $j=1, \ldots, N$ we $\left(\beta_{j} u\right) \circ H_{j} \in W_{2}^{1}\left(Q_{+}\right)$with support in $Q$. Define $\Lambda: W_{2}^{1}(\Omega) \rightarrow W_{2}^{1}(Q) \times\left[W_{2}^{1}(\Omega)\right]^{N}$ by

$$
\Lambda u=\left(\beta_{0} u, \beta_{1} u \circ H^{1}, \ldots, \beta_{N} u \circ H^{N}\right)
$$

Note that we can write $\beta_{0} u \in W_{2}^{\mathrm{o}}(Q)$ as $\beta_{0} u$ has compact support in $\Omega$ and thus, by Friedrichs lemma, it can be approximated by $C_{0}^{\infty}(Q)$ functions. The mapping $\Lambda$ is a linear injection as if $u(x) \neq 0$, then at least one entry of $\Lambda$ must be nonzero as $\beta$ s sum up to 1 . Also, using Proposition 1.48, we can show that the norm on $\Lambda W_{2}^{1}(\Omega)$ is equivalent to the norm on $W_{2}^{1}(\Omega)$ and thus $\Lambda$ is an isomorphism of $W_{2}^{1}(\Omega)$ onto its closed image.

### 1.3.9 Extension operator

We observed that one of the main obstacles in proving that $W_{2}^{1}(\Omega)$ can be obtained by closure of restrictions of $C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ functions to $\Omega$ is that we have no control over the regularization at points close to the boundary of $\Omega$. A remedy could be if we are able to show that any function $W_{2}^{1}(\Omega)$ can be extended to a function from $W_{2}^{1}(\Omega)$.

Indeed, we have
Theorem 1.50. Suppose that $\Omega$ is bounded with a $C^{1}$ boundary $\partial \Omega$. Then there exists a linear extension operator

$$
E: W_{2}^{1}(\Omega) \rightarrow W_{2}^{1}\left(\mathbb{R}^{n}\right)
$$

such that for any $u \in W_{2}^{1}(\Omega)$

1. $\left.E u\right|_{\Omega}=u$;
2. $\|E u\|_{0, \mathbb{R}^{n}} \leq C\|u\|_{0, \Omega}$;
3. $\|E u\|_{1, \mathbb{R}^{n}} \leq C\|u\|_{1, \Omega}$;

Proof. We begin by showing that we can construct an extension operator from $W_{2}^{1}\left(Q_{+}\right)$to $W_{2}^{1}(Q)$. Let $u \in W_{2}^{1}\left(Q_{+}\right)$and define extension by reflection

$$
u^{*}\left(\mathbf{x}^{\prime}, x_{n}\right)= \begin{cases}u\left(\mathbf{x}^{\prime}, x_{n}\right) & \text { for } x_{n}>0 \\ u\left(\mathbf{x}^{\prime},-x_{n}\right) & \text { for } x_{n}<0\end{cases}
$$

where $\mathbf{x}^{\prime}=\left(x_{1}, \ldots, x_{n-1}\right)$. In the same way, we define the odd reflection

$$
u^{\bullet}\left(\mathbf{x}^{\prime}, x_{n}\right)= \begin{cases}u\left(\mathbf{x}^{\prime}, x_{n}\right) & \text { for } x_{n}>0 \\ -u\left(\mathbf{x}^{\prime},-x_{n}\right) & \text { for } x_{n}<0\end{cases}
$$

Further, we define a cut-off function close to $x_{n}=0$, that is, we take a $C^{\infty}(\mathbb{R})$ function $\eta$ which satisfies $\eta(t)=1$ for $t \geq 1$ and $\eta(t)=0$ for $t \leq 1 / 2$ and define $\eta_{k}\left(x_{n}\right)=\eta\left(k x_{n}\right)$. Let us take $\phi \in C_{0}^{\infty}(Q)$ and consider, for $1 \leq i \leq n-1$,

$$
\int_{Q} u^{*} \partial_{x_{i}} \phi d \mathbf{x}=\int_{Q_{+}} u \partial_{x_{i}} \psi d \mathbf{x}
$$

where $\psi\left(\mathbf{x}^{\prime}, x_{n}\right)=\phi\left(\mathbf{x}^{\prime}, x_{n}\right)+\phi\left(\mathbf{x}^{\prime},-x_{n}\right)$. Typically, $\psi$ is not zero at $Q_{0}$ and cannot be used as a test function. However, $\eta_{k}\left(x_{n}\right) \psi(\mathbf{x}) \in C_{0}^{\infty}\left(Q_{+}\right)$and we can write

$$
\int_{Q_{+}} u \partial_{x_{i}}\left(\eta_{k} \psi\right) d \mathbf{x}=\int_{Q_{+}}\left(\partial_{x_{i}} u\right) \eta_{k} \psi d \mathbf{x}
$$

However, $\partial_{x_{i}}\left(\eta_{k} \psi\right)=\eta_{k} \partial_{x_{i}} \psi$ as $\eta$ does not depend on $x_{i}, i=1, \ldots, n-1$ and hence

$$
\int_{Q_{+}} \eta_{k} u \partial_{x_{i}} \psi d \mathbf{x}=-\int_{Q_{+}}\left(\partial_{x_{i}} u\right) \eta_{k} \psi d \mathbf{x} .
$$

We can pass to the limit by dominated convergence getting

$$
\int_{Q_{+}} u \partial_{x_{i}} \psi d \mathbf{x}=-\int_{Q_{+}}\left(\partial_{x_{i}} u\right) \psi d \mathbf{x}
$$

so that, returning to $Q$

$$
\int_{Q} u^{*} \partial_{x_{i}} \phi d \mathbf{x}=-\int_{Q_{+}}\left(\partial_{x_{i}} u\right) \psi d \mathbf{x}=-\int_{Q}\left(\partial_{x_{i}} u\right)^{*} \phi d \mathbf{x} .
$$

Now let us consider differentiability with respect to $x_{n}$. Again, taking $\phi \in$ $C_{0}^{\infty}(Q)$

$$
\int_{Q} u^{*} \partial_{x_{n}} \phi d \mathbf{x}=\int_{Q_{+}} u \partial_{x_{n}} \chi d \mathbf{x}
$$

where $\chi\left(\mathbf{x}^{\prime}, x_{n}\right)=\phi\left(\mathbf{x}^{\prime}, x_{n}\right)-\phi\left(\mathbf{x}^{\prime},-x_{n}\right)$. If we again use $\eta_{k}$, then

$$
\partial_{x_{n}}\left(\eta_{k} \chi\right)=\eta_{k} \partial_{x_{n}} \chi+\chi \partial_{x_{n}} \eta_{k}
$$

where $\partial_{x_{n}} \eta_{k}\left(x_{n}\right)=k \eta^{\prime}\left(k x_{n}\right)$. Then

$$
\begin{aligned}
k\left|\int_{Q_{+}} u(\mathbf{x}) \eta^{\prime}\left(k x_{n}\right) \chi(\mathbf{x}) d x\right| & \leq k C M \int_{Q_{0}}\left(\int_{0}^{1 / k}|u(\mathbf{x})| x_{n} d x_{n}\right) d \mathbf{x}^{\prime} \\
& \leq C M \int_{Q_{+}}|u(\mathbf{x})| d \mathbf{x} \rightarrow 0
\end{aligned}
$$

as $k \rightarrow \infty$, where $C=\sup _{t \in[0,1]}\left|\eta^{\prime}(t)\right|$ and $M$ is obtained from the estimate

$$
\left|\chi\left(\mathbf{x}^{\prime}, x_{n}\right) \leq M\right| x_{n} \mid
$$

on $Q$. Thus

$$
\int_{Q_{+}} u \partial_{x_{n}} \eta_{k} \chi d \mathbf{x}=\int_{Q_{+}} u\left(\eta_{k} \partial_{x_{n}} \chi+\chi \partial_{x_{n}} \eta_{k}\right) d \mathbf{x} \rightarrow \int_{Q_{+}} u \eta_{k} \partial_{x_{n}} \chi
$$

and thus we obtain in the limit

$$
\int_{Q_{+}} u \partial_{x_{n}} \chi d \mathbf{x}=-\int_{Q_{+}}\left(\partial_{x_{n}}\right) u \chi d \mathbf{x}
$$

Returning to $Q$, we obtain

$$
\int_{Q} u^{*} \partial_{x_{n}} \phi d \mathbf{x}=\int_{Q_{+}} u \partial_{x_{n}} \chi d \mathbf{x}=\int_{Q}\left(\partial_{x_{n}} u\right)^{\bullet} \phi d \mathbf{x}
$$

We also obtain estimates

$$
\left\|u^{*}\right\|_{0, Q} \leq 2\|u\|_{0, Q_{+}} \quad\left\|u^{*}\right\|_{1, Q} \leq 2\|u\|_{1, Q_{+}}
$$

Now we can pass to the general result. Let $u \in W_{2}^{1}(\Omega), \Omega$ bounded with $C^{1}$ boundary. Let $\left\{B_{j}, H^{j}\right\}_{j=1}^{N}$ be the atlas on the boundary and $\left\{G_{j}\right\}_{j=1}^{N}$ be the finite subcover constructed in the previous section, that is $G_{0} \subset \bar{G}_{0} \subset \Omega$, $\bar{G}_{j} \subset B_{j}$ with $\partial \Omega \subset \bigcup G_{j}$ and let $\{\beta\}_{j=1}^{N}$ be a subordinate partition of unity. Then we take

$$
u=\sum_{j=0}^{N} \beta_{j} u=\sum_{j=0}^{N} u_{j}
$$

with $u_{0} \in{ }_{W}^{\mathrm{o}}{ }_{2}^{1}(\Omega)$ and $u_{j} \in W_{2}^{1}\left(\Omega \cap B_{j}\right)$. Clearly, $\left\|u_{0}\right\|_{1, \Omega} \leq C_{0}\|u\|_{1, \Omega}$ and $\left\|u_{j}\right\|_{1, \Omega \cap B_{j}} \leq C_{j}\|u\|_{1, \Omega}, j=1, \ldots, n$. The function $u_{0}$ can be extended to $\hat{u}_{0} \in W_{2}^{1}\left(\mathbb{R}^{n}\right)$ by zero in a continuous way. Then $v_{j}:=u_{j} \circ H^{j} \in W_{2}^{1}\left(Q_{+}\right)$ and we can extend by reflection to $v_{j}^{*} \in W_{2}^{1}(Q)$. We note that $v_{j}^{*}$ has support in $Q$ since the support of $u_{j}$ only can touch $\partial\left(B_{j} \cap \Omega\right.$ at the points of $\partial \Omega$. Again,

$$
\left\|v_{j}^{*}\right\|_{1, Q} \leq 2\left\|v_{j}\right\|_{1, Q_{+}} \leq C_{j}^{\prime \prime}\left\|u_{j}\right\|_{1, \Omega \cap B_{j}} \leq C_{j}^{\prime}\|u\|_{1, \Omega}
$$

Next, we define $w_{j}=v_{j}^{*} \circ\left(H^{j}\right)^{-1} \in W_{2}^{1}\left(B_{j}\right)$, again with $\left\|w_{j}\right\|_{1, B_{j}} \leq C_{j}^{\prime \prime}\|u\|_{1, \Omega}$. Moreover, we have $w_{j}(\mathbf{x})=u_{j}(\mathbf{x})$ whenever $\mathbf{x} \in B_{j} \cap \bar{\Omega}$ as

$$
v_{j}^{*}\left(\left(H^{j}\right)^{-1}(\mathbf{x})\right)=v_{j}\left(\left(H^{j}\right)^{-1}(\mathbf{x})\right)=u_{j}\left(H^{j}\left(\left(H^{j}\right)^{-1}(\mathbf{x})\right)\right)=u_{j}(\mathbf{x})
$$

for such $\mathbf{x}$. We also notice that for each $j=1, \ldots, N$, support of $w_{j}$ is contained in $B_{j}$ and thus can extend $w_{j}$ by zero to $\mathbb{R}^{n}$ continuously in $W_{2}^{1}\left(\mathbb{R}^{n}\right)$
and denote this extension by $\hat{u}_{j}$. We note that $\hat{u}_{j}(\mathbf{x})=u_{j}(\mathbf{x})$ for $\mathbf{x} \in \bar{\Omega}$. Indeed, if $\mathbf{x} \in \bar{\Omega}$, for a given $j$ either $\mathbf{x} \in B_{j} \cap \bar{\Omega}$ and then $\hat{u}_{j}(\mathbf{x})=w_{j}(\mathbf{x})=u_{j}(\mathbf{x})$ or $\mathbf{x} \notin B_{j} \cap \bar{\Omega}$ in which case $\hat{u}_{j}(\mathbf{x})=0$ but then also $u_{j}(\mathbf{x})=0$ by definition. The same argument applies to $j=0$. Now we define the operator

$$
E u=\hat{u}_{0}+\sum_{j=1}^{n} \hat{u}_{j}
$$

and we clearly have

$$
E u(\mathbf{x})=\hat{u}_{0}(\mathbf{x})+\sum_{j=1}^{n} \hat{u}_{j}(\mathbf{x})=u_{0}(\mathbf{x})+\sum_{j=1}^{n} u_{j}(\mathbf{x})=u(\mathbf{x})
$$

Linearity and continuity follows from continuity and linearity of each operation and the fact that the sum is finite.

Remark 1.51. Similar argument allows to prove that there is an extension from $W_{2}^{m}(\Omega)$ to $W_{2}^{m}\left(\mathbb{R}^{n}\right)$ (as well as for $W_{p}^{m}(\Omega), 1 \leq p \leq \infty$ ) but this requires the boundary to be a $C^{m}$-manifold (so that the flattening preserves the differentiability). However, the extension across the hyperplane $x_{n}=0$ is done according to the following reflection
$u^{*}\left(\mathbf{x}^{\prime}, x_{n}\right)= \begin{cases}u\left(\mathbf{x}^{\prime}, x_{n}\right) & \text { for } x_{n}>0 \\ \lambda_{1} u\left(\mathbf{x}^{\prime},-x_{n}\right)+\lambda_{2} u\left(\mathbf{x}^{\prime},-\frac{x_{n}}{2}\right)+\ldots+\lambda_{m} u\left(\mathbf{x}^{\prime},-\frac{x_{n}}{m}\right) & \text { for } x_{n}<0,\end{cases}$
where $\lambda_{1}, \ldots, \lambda_{m}$ is the solution of the system

$$
\begin{aligned}
\lambda_{1}+\lambda_{2}+\ldots+\lambda_{m} & =1, \\
-\left(\lambda_{1}+\lambda_{2} / 2+\ldots+\lambda_{m} / m\right) & =1 \\
& \cdots \\
(-1)^{m}\left(\lambda_{1}+\lambda_{2} / 2^{m-1}+\ldots+\lambda_{m} / m^{m-1}\right) & =1
\end{aligned}
$$

These conditions ensure that the derivatives in the $x_{n}$ direction are continuous $\operatorname{across} x_{n}=0$.

An immediate consequence of the extension theorem is
Theorem 1.52. Let $\Omega$ be a bounded set with a $C^{1}$ boundary $\partial \Omega$ and $u \in$ $W_{2}^{1}(\Omega)$. Then there exits $\left(u_{n}\right)_{n \in \mathbb{N}}, u_{n} \in C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ such that

$$
\left.\lim _{n \rightarrow \infty} u_{n}\right|_{\Omega}=u, \quad \text { in } W_{2}^{1}(\Omega)
$$

In other words, the set of restriction to $\Omega$ of functions from $C_{0}^{\infty}(\Omega)$ is dense in $W_{2}^{1}(\Omega)$.

Proof. If $\Omega$ is bounded then, using Theorem 1.50, we can extend $u$ to a function $E u \in W_{2}^{1}\left(\mathbb{R}^{n}\right)$ with bounded support. The existence of a $C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ sequence converging to $u$ follows from the Friedrichs lemma. If $\Omega$ is unbounded (but not equal to $\mathbb{R}^{n}$ ), then first we approximate $u$ by a sequence $\left(\chi_{n} u\right)_{n \in \mathbb{N}}$ where $\chi_{n}$ are cut-off functions. Next we construct an extension of $\chi_{n} u$ to $\mathbb{R}^{n}$. This is possible as it involves only the part of $\partial \Omega$ intersecting the ball $B(0,2 n+1)$ and $\chi_{n}$ is equal to zero where the sphere intersects $\partial \Omega$. For this extension we pick up an approximating function from $C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$.

### 1.4 Basic applications of the density theorem

### 1.4.1 Sobolev embedding

In Subsection 1.1.2 we have seen that in one dimension it is possible to identify a $W_{2}^{1}(\mathbb{R})$ function. Unfortunately, this is not true in higher dimensions.

Example 1.53. We can consider in $D=\left\{(x, y) \in \mathbb{R}^{2} ; x^{2}+y^{2}<1\right\}$

$$
u(x, y)=\left|\frac{1}{2} \ln \left(x^{2}+y^{2}\right)\right|^{1 / 3}=(-\ln r)^{1 / 3}
$$

The function $u$ is not continuous (even not bounded) at $(x, y)=(0,0)$. It is in $L_{2}(D)$ and for derivatives we have

$$
u_{x}=-\frac{1}{3}(-\ln r)^{-2 / 3} \frac{x}{r^{2}}, \quad u_{y}=-\frac{1}{3}(-\ln r)^{-2 / 3} \frac{y}{r^{2}}
$$

and, since

$$
\int_{D}\left(u_{x}^{2}+u_{y}^{2}\right) d x d y=\frac{2}{9} \int_{0}^{1} \frac{d r}{r(-\ln r)^{4 / 3}}=\frac{2}{9} \int_{1}^{\infty} u^{-4 / 3} d u<\infty
$$

we see that $u \in W_{2}^{1}(D)$.
However, there is still a link between Sobolev spaces and classical calculus provided we take sufficiently high order of derivatives (or index $p$ in $L_{p}$ spaces). The link is provided by the Sobolev lemma.

Let $\Omega$ be an open and bounded subset of $\mathbb{R}^{n}$. We say that $\Omega$ satisfies the cone condition if there are numbers $\rho>0$ and $\gamma>0$ such that each $\mathbf{x} \in \Omega$ is a vertex of a cone $K(\mathbf{x})$ of radius $\rho$ and volume $\gamma \rho^{n}$. Precisely speaking, if $\sigma_{n}$ is the $n-1$ dimensional measure of the unit sphere in $\mathbb{R}^{n}$, then the volume of a ball of radius $\rho$ is $\sigma_{n} \rho^{n} / n$ and then the (solid) angle of the cone is $\gamma n / \omega_{n}$.
Lemma 1.54. If $\Omega$ satisfies the cone condition, then there exists a constant $C$ such that for any $u \in C^{m}(\bar{\Omega})$ with $2 m>n$ we have

$$
\begin{equation*}
\sup _{\mathbf{x} \in \Omega}|u(\mathbf{x})| \leq C\|u\|_{m} \tag{1.62}
\end{equation*}
$$

Proof. Let us introduce a cut-off function $\phi \in C_{0}^{\infty}(\mathbb{R})$ which satisfies $\phi(t)=1$ for $|t| \leq 1 / 2$ and $\phi(t)=0$ for $|t| \geq 1$. Define $\tau(t)=\phi(t / \rho)$ and note that there are constants $A_{k}, k=1,2, \ldots$ such that

$$
\begin{equation*}
\left|\frac{d^{k} \tau(t)}{d t^{k}}\right| \leq \frac{A_{k}}{\rho^{k}} \tag{1.63}
\end{equation*}
$$

Let us take $u \in C^{m}(\bar{\Omega})$ and assume $2 m>n$. For $\mathbf{x} \in \bar{\Omega}$ and the cone $K(\mathbf{x})$ we integrate along the ray $\{\mathbf{x}+r \boldsymbol{\omega} ; 0 \leq r \leq \rho,|\boldsymbol{\omega}|=1$

$$
u(\mathbf{x})=-\int_{0}^{\rho} D_{r}(\tau(r) u(\mathbf{x}+r \boldsymbol{\omega})) d r
$$

Integrating over the surface $\Gamma$ of the cone we get

$$
\int_{\Gamma} \int_{0}^{\rho} D_{r}(\tau(r) u(\mathbf{x}+\boldsymbol{\omega})) d r d \boldsymbol{\omega}=-u(\mathbf{x}) \int_{C} d \boldsymbol{\omega}=-u(\mathbf{x}) \frac{\gamma n}{\omega_{n}}
$$

Next we integrate $m-1$ times by parts, getting

$$
u(\mathbf{x})=\frac{(-1)^{m} \omega_{n}}{\gamma n(m-1)!} \int_{C} \int_{0}^{\rho} D_{r}^{m}(\tau(r) u(\mathbf{x}+r \boldsymbol{\omega})) r^{m-1} d r d \boldsymbol{\omega} .
$$

and changing to Cartesian coordinates and applying Cauchy-Schwarz inequality we obtain

$$
\begin{aligned}
|u(\mathbf{x})|^{2} & \leq\left(\frac{\omega_{n}}{\gamma n(m-1)!} \int_{K(\mathbf{x})}\left|D_{r}^{m}(\tau u)\right| r^{m-n} d \mathbf{y}\right)^{2} \\
& \leq\left(\frac{\omega_{n}}{\gamma n(m-1)!}\right)^{2} \int_{K(\mathbf{x})}\left|D_{r}^{m}(\tau u)\right|^{2} d \mathbf{y} d \mathbf{y} \int_{K(\mathbf{x})} r^{2(m-n)} d \mathbf{y}
\end{aligned}
$$

The last term can be evaluated as

$$
\int_{K(\mathbf{x})} r^{2(m-n)} d \mathbf{y}=\int_{C} \int_{0}^{\rho} r^{2 m-n-1} d r d \boldsymbol{\omega}=\frac{\gamma n \rho^{2 m-n}}{\omega_{n}(2 m-n)}
$$

so that

$$
\begin{equation*}
|u(\mathbf{x})|^{2} \leq C(m, n) \rho^{2 m-n} \int_{K(\mathbf{x})}\left|D_{r}^{m}(\tau u)\right|^{2} d \mathbf{y} \tag{1.64}
\end{equation*}
$$

Let us estimate the derivative. From (1.63) we obtain by the chain rule and the Leibniz formula

$$
\left|D_{r}^{m}(\tau u)\right|=\left|\sum_{k=0}^{m}\binom{n}{k} D_{r}^{m-k} \tau D_{r}^{k} u\right| \leq \sum_{k=0}^{m}\binom{n}{k} \frac{A_{m-k}}{\rho^{m-k}}\left|D_{r}^{k} u\right|
$$

hence

$$
\left|D_{r}^{m}(\tau u)\right|^{2} \leq C^{\prime} \sum_{k=0}^{m} \frac{1}{\rho^{2(m-k)}}\left|D_{r}^{k} u\right|^{2}
$$

for some constant $C^{\prime}$. With this estimate we can re-write (1.64) as

$$
\begin{equation*}
|u(\mathbf{x})|^{2} \leq C(m, n) C^{\prime} \sum_{k=0}^{m} \rho^{2 k-n} \int_{K(\mathbf{x})}\left|D_{r}^{m}(u)\right|^{2} d \mathbf{y} \tag{1.65}
\end{equation*}
$$

Since by the chain rule

$$
\left|D_{r}^{m} u\right|^{2} \leq C^{\prime \prime} \sum_{|\alpha| \leq k}\left|D^{\alpha} u\right|^{2}
$$

by extending the integral to $\Omega$ we obtain

$$
\sup _{\mathbf{x} \in \Omega}|u(\mathbf{x})| \leq C\|u\|_{m}
$$

which is (1.62).
Theorem 1.55. Assume that $\Omega$ is a bounded open set with $C^{m}$ boundary and let $m>k+n / 2$ where $m$ and $k$ are integers. Then the embedding

$$
W_{2}^{m}(\Omega) \subset C^{k}(\bar{\Omega})
$$

is continuous.
Proof. Under the assumptions, the problem can be reduced to the set $G_{0} \Subset \Omega$ consisting of internal point, separated from the boundary by a fixed positive distance, and points in the boundary strip, covered by sets $\bar{\Omega} \cap B_{j}$ which are transformed onto $Q_{+} \cup Q_{0}$. Any point in $G_{0}$ satisfies the cone conditions. Points on $Q_{0} \cup Q_{+}$also satisfy the condition so, if $u \in W_{2}^{m}(\Omega)$, then extending the boundary components of $\Lambda u$ to $Q$ we obtain functions in $W_{2}^{1}(\Omega)$ and $W_{2}^{1}(Q)$ with compact supports in respective domains. By Friedrichs lemma, restrictions to $\Omega$ and $Q$ of $C^{\infty}\left(\mathbb{R}^{n}\right)$ functions are dense in, respectively, $W_{2}^{m}(\Omega)$ and $W_{2}^{m}(Q)$ and therefore the estimate (1.62) can be extended by density to $W_{2}^{m}(\Omega)$ showing that the canonical injection into $C(\bar{\Omega})$ is continuous. To obtain the result for higher derivatives we substitute higher derivatives of $u$ for $u$ in (1.62). Thus, all components of $\Lambda u$ are they are $C^{k}$ functions. Transferring them back, we see that $u \in C^{k}(\bar{\Omega})$, by regularity of the local atlas and $m>k$, we obtain the thesis.

### 1.4.2 Compact embedding and Rellich-Kondraschov theorem

Lemma 1.56. let $Q=\left\{\mathbf{x} ; a_{j} \leq x_{j} \leq b_{j}\right\}$ be a cube in $\mathbb{R}^{n}$ with edges of length $d>0$. If $u \in C^{1}(\bar{Q})$, then

$$
\begin{equation*}
\|u\|_{0, Q}^{2} \leq d^{-n}\left(\int_{Q} u d \mathbf{x}\right)^{2}+\frac{n d^{2}}{2} \sum_{j=1}^{n}\left\|\partial_{x_{j}} u\right\|_{0, Q}^{2} \tag{1.66}
\end{equation*}
$$

Proof. For any $\mathbf{x}, \mathbf{y} \in Q$ we can write

$$
u(\mathbf{x})-u(\mathbf{y})=\sum_{j=1}^{n} \int_{y_{j}}^{x_{j}} \partial_{x_{j}} u\left(y_{1}, \ldots, y_{j-1}, s, x_{j+1}, \ldots, x_{n}\right) d s
$$

Squaring this identity and using Cauchy-Schwarz inequality we obtain
$u^{2}(\mathbf{x})+u^{2}(\mathbf{y})-2 u(\mathbf{x}) u(\mathbf{y}) \leq n d \sum_{j=1}^{n} \int_{a_{j}}^{b_{j}}\left(\partial_{j} u\right)^{2}\left(y_{1}, \ldots, y_{j-1}, s, x_{j+1}, \ldots, x_{n}\right) d s$.
Integrating the above inequality with respect to all variables, we obtain

$$
2 d^{n}\|u\|_{0, Q}^{2} \leq 2\left(\int_{Q} u d \mathbf{x}\right)^{2}+n d^{n+2} \sum_{j=1}^{n}\left\|\partial_{j} u\right\|_{0, Q}^{2}
$$

as required.
Theorem 1.57. Let $\Omega$ be open and bounded. If the sequence $\left(u_{k}\right)_{k \in \mathbb{N}}$ of elements of ${ }_{W}^{\mathbf{o}}{ }_{2}^{1}(\Omega)$ is bounded, then there is a subsequence which converges in in $L_{2}(\Omega)$. In other words, the injection $\stackrel{\circ}{W}_{2}^{1}(\Omega) \subset L_{2}(\Omega)$ is compact.

Proof. By density, we may assume $u_{k} \in C_{0}^{\infty}$. Let $M=\sup _{k}\left\{\left\|u_{k}\right\|_{1}\right\}$. We enclose $\Omega$ in a cube $Q$; we may assume the edges of $Q$ to be of unit length. Further, we extend each $u_{k}$ by zero to $Q \backslash \Omega$.

We decompose $Q$ into $N^{n}$ cubes of edges of length $1 / N$. Since clearly $\left(u_{k}\right)_{k \in \mathbb{N}}$ is bounded in $L_{2}(Q)$ it contains a weakly convergent subsequence (which we denote again by $\left(u_{k}\right)_{k \in \mathbb{N}}$ ). For any $\epsilon^{\prime}$ there is $n_{0}$ such that

$$
\begin{equation*}
\left|\int_{Q_{j}}\left(u_{k}-u_{l}\right) d \mathbf{x}\right|<\epsilon^{\prime}, \quad k, l \geq n_{0} \tag{1.67}
\end{equation*}
$$

for each $j=1, \ldots, N^{n}$. Now, we apply (1.66) on each $Q_{j}$ and sum over all $j$ getting

$$
\left\|u_{k}-u_{l}\right\|_{0, Q}^{2} \leq N^{n} \epsilon^{\prime}+\frac{n}{2 N^{2}} 2 M^{2}
$$

Now, we see that for a fixed $\epsilon$ we can find $N$ large that $n M^{2} / N^{2}<e$ and, having fixed $N$, for $\epsilon^{\prime}=\epsilon / 2 N^{n}$ we can find $n_{0}$ such that (1.67) holds. Thus $\left(u_{k}\right)_{k \in \mathbb{N}}$ is Cauchy in $L_{2}(\Omega)$.

Corollary 1.58. If $\Omega$ is a bounded open subset of $\mathbb{R}^{n}$, then the embedding $\stackrel{\circ}{W}_{2}^{m}(\Omega) \subset \stackrel{\circ}{W}_{2}^{m-1}(\Omega)$ is compact.

Proof. Applying the previous theorem to the sequences of derivatives, we see that the derivatives form bounded sequences in $W_{2}^{1}(\Omega)$ and thus contain subsequences converging in $L_{2}(\Omega)$. Selecting common subsequence we get convergence in $W_{2}^{1}(\Omega)$ etc, (by closedness of derivatives).

Theorem 1.59. If $\partial \Omega$ is a $C^{m}$ boundary of a bounded open set $\Omega$. Then the embedding $W_{2}^{m}(\Omega) \subset W_{2}^{m-1}(\Omega)$ is compact.

Proof. The result follows by extension to $\stackrel{\circ}{W}_{2}^{m}\left(\Omega^{\prime}\right)$ where $\Omega^{\prime}$ is a bounded set containing $\Omega$.

### 1.4.3 Trace theorems

We know that if $u \in W_{2}^{m}(\Omega)$ with $m>n / 2$ then $u$ can be represented by a continuous function and thus can be assigned a value at the boundary of $\Omega$ (or, in fact, at any point). The requirement on $m$ is, however, too restrictive - we have solved the Dirichlet problem, which requires a boundary value of the solution, in $\stackrel{\circ}{W}_{2}^{1}(\Omega)$. In this space, unless $n=1$, the solution need not be continuous. It turns out that it is possible to give a meaning to the operation of taking the boundary value of a function even if it is not continuous.

We begin with the simplest (nontrivial) case when $\Omega=\mathbb{R}_{+}^{n}:=\{\mathbf{x} ; \mathbf{x}=$ $\left.\left(\mathbf{x}^{\prime}, x_{n}\right), 0<x_{n}\right\}$.
Theorem 1.60. The trace operator $\gamma_{0}: C^{1}\left(\overline{\mathbb{R}_{+}^{n}}\right) \cap W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right) \rightarrow C^{0}\left(\mathbb{R}^{n-1}\right)$ defined by

$$
\left(\gamma_{0} \phi\right)\left(\mathbf{x}^{\prime}\right)=\phi\left(\mathbf{x}^{\prime}, 0\right), \quad \phi \in C^{1}\left(\overline{\mathbb{R}_{+}^{n}}\right) \cap W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right), \mathbf{x}^{\prime} \in \mathbb{R}^{n-1}
$$

has a unique extension to a continuous linear operator $\gamma_{0}: W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right) \rightarrow$ $L_{2}\left(\mathbb{R}^{n-1}\right)$ whose range in dense in $L_{2}\left(\mathbb{R}^{n-1}\right)$. The extension satisfies

$$
\gamma_{0}(\beta u)=\gamma_{0}(\beta) \gamma_{0}(u), \quad \beta \in C^{1}\left(\overline{\mathbb{R}_{+}^{n}}\right) \cap L_{\infty}\left(\mathbb{R}_{+}^{n}\right), u \in W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)
$$

Proof. Let $\phi \in C^{1}\left(\overline{\mathbb{R}_{+}^{n}}\right) \cap W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$. Then, from continuity, for any $\mathbf{x}^{\prime}$, $\partial_{x_{n}}\left|\phi\left(\mathbf{x}^{\prime}, x_{n}\right)\right|^{2} \in L_{2}\left(\mathbb{R}_{+}\right)$we can write

$$
\left|\phi\left(\mathbf{x}^{\prime}, r\right)\right|^{2}-\left|\phi\left(\mathbf{x}^{\prime}, 0\right)\right|^{2}=\int_{0}^{r} \partial_{x_{n}}\left|u\left(\mathbf{x}^{\prime}, x_{n}\right)\right|^{2} d x_{n}
$$

and thus $\left|\phi\left(\mathbf{x}^{\prime}, r\right)\right|^{2}$ has a limit which must equal 0 . Hence

$$
\left|\phi\left(\mathbf{x}^{\prime}, 0\right)\right|^{2}=-\int_{0}^{\infty} \partial_{x_{n}}\left|\phi\left(\mathbf{x}^{\prime}, x_{n}\right)\right|^{2} d x_{n}
$$

Integrating over $\mathbb{R}^{n-1}$ we obtain

$$
\begin{aligned}
\left\|\phi\left(\mathbf{x}^{\prime}, 0\right)\right\|_{0, \mathbb{R}^{n-1}}^{2} & \leq 2 \int_{\mathbb{R}_{+}^{n}} \partial_{x_{n}} \phi(\mathbf{x}) \phi(\mathbf{x}) d \mathbf{x} \\
& \leq 2\left\|\partial_{x_{n}} \phi\right\|_{0, \mathbb{R}_{+}^{n}}\|\phi\|_{0, \mathbb{R}_{+}^{n}} \leq\left\|\partial_{x_{n}} \phi\right\|_{0, \mathbb{R}_{+}^{n}}^{2}+\|\phi\|_{0, \mathbb{R}_{+}^{n}}^{2}
\end{aligned}
$$

Hence, by density, the operation of taking value at $x_{n}=0$ extends to $W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$.
If $\phi \in C_{0}^{\infty}\left(\mathbb{R}^{n-1}\right)$ and $\tau$ is a truncation function $\tau(t)=1$ for $|t| \leq 1$ and $\tau(t)=0$ for $|t| \geq 0$ then $\phi(\mathbf{x})=\psi\left(\mathbf{x}^{\prime}\right) \tau\left(x_{n}\right) \in C^{1}\left(\overline{\mathbb{R}_{+}^{n}}\right) \cap W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$ and $\gamma_{0}(\phi)=\psi$ so that the range of the trace operator contains $C_{0}^{\infty}\left(\mathbb{R}^{n-1}\right)$ and thus is dense. The last identity follows from continuity of the trace operator and of the operator of multiplication by bounded differentiable functions in $W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$.

Theorem 1.61. Let $u \in W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$. Then $u \in W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$ if an only if $\gamma_{0}(u)=0$,
Proof. If $u \in \stackrel{o}{W}_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$, then $u$ is the limit of a sequence $\left(\phi_{k}\right)_{k \in \mathbb{N}}$ from $C_{0}^{\infty}\left(\mathbb{R}_{+}^{n}\right)$ in $W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$. Since $\gamma_{0}\left(\phi_{k}\right)=0$ for any $k$, we obtain $\gamma_{0}(u)=0$.

Conversely, let $u \in W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$ with $\gamma_{0} u=0$. By using the truncating functions, we may assume that $u$ has compact support in $\overline{\mathbb{R}_{+}^{n}}$.

Next we use the truncating functions $\eta_{k} \in C^{\infty}(\mathbb{R})$, as in Theorem 1.50, by taking function $\eta$ which satisfies $\eta(t)=1$ for $t \geq 1$ and $\eta(t)=0$ for $t \leq 1 / 2$ and define $\eta_{k}\left(x_{n}\right)=\eta\left(k x_{n}\right)$. To simplify notation, we assume that $0 \leq \eta^{\prime} \leq 3$ for $t \in[1 / 2,1]$ so that $0 \leq \eta_{k}^{\prime}\left(x_{n}\right) \leq 3 k$. Then the extension by 0 to $\mathbb{R}_{-}^{n}$ of $\mathbf{x} \rightarrow \eta_{k}\left(x_{n}\right) u\left(\mathbf{x}^{\prime}, x_{n}\right)$ is in $W_{2}^{1}\left(\mathbb{R}^{n}\right)$ and can be approximated by $C_{0}^{\infty}\left(\mathbb{R}_{+}^{n}\right)$ functions in $W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$. Hence, we have to prove that $\eta_{k} u \rightarrow u$ in $W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$.

As in the proof of Theorem 1.50 we can prove $\eta_{k} u \rightarrow u$ in $L_{2}\left(\mathbb{R}_{+}^{n}\right)$ and for each $i=1, \ldots, n-1, \partial_{x_{i}}\left(\eta_{k} u\right)=\eta_{k} \partial_{x_{i}} u \rightarrow \partial_{x_{i}} u$ in $L_{2}\left(\mathbb{R}_{+}^{n}\right)$ as $k \rightarrow \infty$.

Since

$$
\partial_{x_{n}}\left(\eta_{k} u\right)=u \partial_{x_{n}} \eta_{k}+\eta_{k} \partial_{x_{n}} u
$$

we see that we have to prove that $u \partial_{x_{n}} \eta_{k} \rightarrow 0$ in $L_{2}\left(\mathbb{R}_{+}^{n}\right)$ as $k \rightarrow \infty$. For this, first we prove that if $\gamma_{0}(u)=0$, then

$$
\begin{equation*}
u\left(\mathbf{x}^{\prime}, s\right)=\int_{0}^{s} \partial_{x_{n}} u\left(\mathbf{x}^{\prime}, t\right) d t \tag{1.68}
\end{equation*}
$$

almost everywhere on $\mathbb{R}_{+}^{n}$. Indeed, let $u_{r}$ be a bounded support $C^{1}$ function approximating $u$ in $W_{2}^{1}\left(\mathbb{R}_{+}^{n}\right)$. Then $\int_{0}^{s} \partial_{x_{n}} u_{r}\left(\mathbf{x}^{\prime}, t\right) d t \rightarrow \int_{0}^{s} \partial_{x_{n}} u\left(\mathbf{x}^{\prime}, t\right) d t$ in
$L_{2}\left(\mathbb{R}_{+}^{n}\right)$. This follows from $\partial_{x_{n}} u_{r} \rightarrow \partial_{x_{n}} u$ in $L_{2}\left(\mathbb{R}_{+}^{n}\right)$ and, taking $Q$ to be the box enclosing support of all $u_{r}, u$, with edges of length at most $d$

$$
\begin{aligned}
& \int_{Q}\left|\int_{0}^{s} \partial_{x_{n}} u_{r}\left(\mathbf{x}^{\prime}, t\right) d t-\int_{0}^{s} \partial_{x_{n}} u\left(\mathbf{x}^{\prime}, t\right) d t\right|^{2} d \mathbf{x} \\
& \leq d^{2} \int_{Q}\left|\partial_{x_{n}} u_{r}\left(\mathbf{x}^{\prime}, t\right)-\partial_{x_{n}} u\left(\mathbf{x}^{\prime}, t\right)\right|^{2} d \mathbf{x}
\end{aligned}
$$

Then we have, for any $s, 0 \leq s \leq d$
$\int_{Q}\left|\int_{0}^{s} \partial_{x_{n}} u_{r}\left(\mathbf{x}^{\prime}, t\right) d t-u_{r}\left(\mathbf{x}^{\prime}, s\right)\right|^{2} d \mathbf{x}=\int_{Q}\left|u_{r}\left(\mathbf{x}^{\prime}, 0\right)\right|^{2} d \mathbf{x}=d \int_{\mathbb{R}^{n-1}}\left|u_{r}\left(\mathbf{x}^{\prime}, 0\right)\right|^{2} d \mathbf{x}^{\prime}$
and, since the right hand side goes to zero as $r \rightarrow \infty$, we obtain (1.68). Then, by Cauchy-Schwarz inequality

$$
\left|u\left(\mathbf{x}^{\prime}, s\right)\right|^{2} \leq s \int_{0}^{s}\left|\partial_{x_{n}} u\left(x^{\prime}, t\right)\right|^{2} d t
$$

and therefore

$$
\begin{aligned}
& \int_{0}^{\infty}\left|\eta_{k}^{\prime}(s) u\left(\mathbf{x}^{\prime}, s\right)\right|^{2} d s \leq 9 k^{2} \int_{0}^{2 / k} s \int_{0}^{s}\left|\partial_{x_{n}} u\left(\mathbf{x}^{\prime}, t\right)\right|^{2} d t d s \\
& 18 k \int_{0}^{2 / k} \int_{0}^{s}\left|\partial_{x_{n}} u\left(\mathbf{x}^{\prime}, t\right)\right|^{2} d t d s=18 k \int_{0}^{2 / k} \int_{t}^{2 / k}\left|\partial_{x_{n}} u\left(\mathbf{x}^{\prime}, t\right)\right|^{2} d s d t \\
\leq & 36 \int_{0}^{2 / k}\left|\partial_{x_{n}} u\left(\mathbf{x}^{\prime}, t\right)\right|^{2} d t
\end{aligned}
$$

Integration over $\mathbb{R}^{n-1}$ gives

$$
\left\|\eta_{k}^{\prime} u\right\|_{0, \mathbb{R}_{+}^{n}}^{2} \leq 36 \int_{\mathbb{R}^{n-1} \times 2 / k}\left|\partial_{x_{n}} u\right|^{2} d \mathbf{x}
$$

which tends to 0 .
The consideration above can be extended to the case where $\Omega$ is an open bounded region in $\mathbb{R}^{n}$ lying locally on one side of its $C^{1}$ boundary. Using the partition of unity, we define

$$
\gamma_{0}(u):=\sum_{j=1}^{N}\left(\gamma_{0}\left(\left(\beta_{j} u\right) \circ H^{j}\right)\right) \circ\left(H^{j}\right)^{-1}
$$

It is clear that if $u \in C^{1}(\bar{\Omega})$, then $\gamma_{0} u$ is the restriction of $u$ to $\partial \Omega$. Thus, we have the following result

Theorem 1.62. Let $\Omega$ be a bounded open subset of $\mathbb{R}^{n}$ which lies on one side of its boundary $\partial \Omega$ which is assumed to be a $C^{1}$ manifold. Then there exists a unique continuous and linear operator $\gamma_{0}: W_{2}^{1}(\Omega) \rightarrow L_{2}(\partial \Omega)$ such that for each $u \in C^{1}(\bar{\Omega}), \gamma_{0}$ is the restriction of $u$ to $\partial \Omega$. The kernel of $\gamma_{0}$ is equal to $\stackrel{\circ}{W_{2}^{1}}(\Omega)$ and its range is dense in $L_{2}(\partial \Omega)$.

### 1.4.4 Regularity of variational solutions to the Dirichlet problem

From Subsection 1.3 .6 we know that there is a unique variational solution $u \in \stackrel{o}{W}_{2}^{1}(\Omega)$ of the problem

$$
\int_{\Omega} \nabla u \nabla v d \mathbf{x}=<f, v>_{\left(W_{2}^{\mathrm{o}}(\Omega)\right)^{*} \times W_{2}^{\mathrm{o}}(\Omega)}, \quad v \in W_{2}^{\mathrm{o}}(\Omega) .
$$

Moreover, now we can say that $\gamma_{0} u=0$ on $\partial \Omega$ (provided $\partial \Omega$ is $C^{1}$ ).
We have the following theorem
Theorem 1.63. Let $\Omega \subset \mathbb{R}^{n}$ be an open bounded set with $C^{2}$ boundary (or $\Omega=\mathbb{R}_{+}^{n}$ ). Let $f \in L_{2}(\Omega)$ and let $u \in W_{2}^{1}(\Omega)$ satisfy

$$
\begin{equation*}
\int_{\Omega} \nabla u \nabla v d \mathbf{x}=(f, v), \quad v \in \stackrel{o}{W}_{2}^{1}(\Omega) \tag{1.69}
\end{equation*}
$$

Then $u \in W_{2}^{2}(\Omega)$ and $\|u\|_{2, \Omega} \leq C\|f\|_{0, \Omega}$ where $C$ is a constant depending only on $\Omega$. Furthermore, if $\Omega$ is of class $C^{m+2}$ and $f \in W_{2}^{m}(\Omega)$, then

$$
u \in W_{2}^{m+2}(\Omega) \quad \text { and } \quad\|u\|_{m+2, \Omega} \leq C\|f\|_{m, \Omega}
$$

In particular, if $m \geq n / 2$, then $u \in C^{2}(\bar{\Omega})$ is a classical solution.
Moreover, if $\Omega$ is bounded, then the solution operator $G: L_{2}(\Omega) \rightarrow \stackrel{\circ}{W}_{2}^{1}(\Omega)$ is self-adjoint and compact.

Proof. The proof naturally splits into two cases: interior estimates and boundary estimates. Let $\Omega$ be bounded with at least $C^{1}$ boundary and consider the partition of unity $\left\{\beta_{j}\right\}_{j=0}^{N}$ subordinated to the covering $\left\{G_{j}\right\}_{j=0}^{N}$. For the interior estimates let us consider $u_{0}=\beta_{0} u$ and let $v \in W_{2}^{1}(\Omega)$. Then we can write

$$
\begin{aligned}
\int_{\Omega} \nabla\left(\beta_{0} u\right) \nabla v d \mathbf{x} & =\int_{\Omega} \beta_{0} \nabla u \nabla v d \mathbf{x}+\int_{\Omega} u \nabla \beta_{0} \nabla v d \mathbf{x} \\
& =\int_{\Omega} \nabla u \nabla\left(\beta_{0} v\right) d \mathbf{x}-\int_{\Omega} v \nabla u \nabla \beta_{0} d \mathbf{x}+\int_{\Omega} u \nabla v \nabla \beta_{0} d \mathbf{x} \\
& =\int_{\Omega} \nabla u \nabla\left(\beta_{0} v\right) d \mathbf{x}-\int_{\Omega} v \nabla u \nabla \beta_{0} d \mathbf{x}-\int_{\Omega} \nabla\left(u \nabla \beta_{0}\right) v d \mathbf{x} \\
& =\int_{\Omega} \nabla u \nabla\left(\beta_{0} v\right) d \mathbf{x}-2 \int_{\Omega} v \nabla u \nabla \beta_{0} d \mathbf{x}-\int_{\Omega} u v \Delta \beta_{0} d \mathbf{x} \\
& =\int_{\Omega}\left(f \beta_{0}-\Delta \beta_{0} u-2 \nabla u \nabla \beta_{0}\right) v d \mathbf{x}=\int_{\Omega} F v d \mathbf{x}, \quad v \in W_{2}^{1}(\Omega),
\end{aligned}
$$

where $F \in L_{2}(\Omega)$ and we used $v \in \stackrel{o}{W}_{2}^{1}(\Omega)$ to get

$$
\int_{\Omega} u \nabla v \nabla \beta_{0} d \mathbf{x}=-\int_{\Omega} \nabla\left(u \nabla \beta_{0}\right) v d \mathbf{x}
$$

Hence, the function $w=\beta_{0} u$ is the variational solution to the above problem in $\mathbb{R}^{n}$. Let us define $D_{h} u=|\mathbf{h}|^{-1}\left(\tau_{h} u-u\right)$ and take $v=D_{-h}\left(D_{h} w\right)$. It is possible since $w$ has compact support in $\Omega$ and thus $v \in W_{2}^{1}(\Omega)$ for sufficiently small $\mathbf{h}$. Thus we obtain

$$
\int_{\Omega}\left|\nabla D_{h} w\right|^{2} d \mathbf{x}=\int_{\Omega} F D_{-h}\left(D_{h} w\right) d \mathbf{x}
$$

that is,

$$
\begin{equation*}
\left\|D_{h} w\right\|_{1, \Omega}^{2} \leq\|F\|_{0, \Omega}\left\|D_{-h}\left(D_{h} w\right)\right\|_{0, \Omega} \tag{1.70}
\end{equation*}
$$

On the other hand, from Friedrichs lemma, for any $v \in W_{2}^{1}(\Omega)$ with compact support

$$
\begin{equation*}
\left\|D_{-h} v\right\|_{0, \Omega}^{2} \leq\|\nabla v\|_{0, \Omega} \tag{1.71}
\end{equation*}
$$

Applying this to $v=D_{h} u$, we obtain

$$
\left\|D_{h} w\right\|_{1, \Omega}^{2} \leq\|F\|_{0, \Omega}\left\|\nabla D_{h} w\right\|_{0, \Omega} \leq\|F\|_{0, \Omega}\left\|D_{h} w\right\|_{1, \Omega},
$$

that is,

$$
\left\|D_{h} w\right\|_{1, \Omega} \leq\|F\|_{0, \Omega}
$$

In particular, we obtain

$$
\left\|D_{h} \partial_{x_{i}} w\right\|_{0, \Omega} \leq\|F\|_{0, \Omega}, \quad i=1, \ldots, n
$$

which yields $\partial_{x_{i}} w \in W_{2}^{1}(\Omega)$, that is, $w \in W_{2}^{2}(\Omega)$.
In the next step, we shall move to estimates close to the boundary. Let us fix some some set $B_{j}$ and corresponding function $\beta_{j}, 1 \leq j \leq N$ from the partition of unity and drop the index $j$. Then we have a $C^{2}$ diffeomorphism $H: Q \rightarrow B$ the inverse of which we denote $J=H^{-1}$ so that $H\left(Q_{+}\right)=\Omega \cap B$ and $H\left(Q_{0}\right)=\partial \Omega \cap B$. We denote $\mathbf{x}=H(\mathbf{y}), \mathbf{y} \in Q$ and $\mathbf{y}=J(\mathbf{x})$. As before, we see that $w=\beta u$ is a variational solution to

$$
\begin{equation*}
\int_{\Omega \cap B} \nabla w \nabla v d \mathbf{x}=\int_{\Omega \cap B}(f \beta-u \Delta \beta-2 \nabla u \nabla \beta) v d \mathbf{x}=\int_{\Omega \cap B} g v d \mathbf{x}, \quad v \in W_{2}^{1}(\Omega) \tag{1.72}
\end{equation*}
$$

where the Green's formula

$$
\int_{\Omega \cap B} u \nabla v \nabla \beta_{0} d \mathbf{x}=-\int_{\Omega \cap B} \nabla\left(u \nabla \beta_{0}\right) v d \mathbf{x}
$$

can be justified by noting that the integration is actually carried out over the domain $G \Subset B$ and we can use a function $\chi v$, where $\chi$ is equal to 1 on $G$ and has support in $B$, instead of $v$. Function $\chi v \in W_{2}^{1}(\Omega \cap B)$ (as $v$ can be approximated by $\phi$ compactly supported in $\Omega$ and $\chi v$ can be approximated by $\chi \phi$ compactly supported in $\Omega \cap B)$.

Now we transfer (1.72) to $Q_{+}$. We have $z(\mathbf{y})=w(H(\mathbf{y}))$ for $\mathbf{y} \in Q_{+}$or $w(\mathbf{x})=z(J(\mathbf{x}))$ for $\mathbf{x} \in \Omega \cap B$. Let $\psi \in \stackrel{\circ}{W_{2}^{1}}\left(Q_{+}\right)$and $\phi(\mathbf{x})=\psi(J(\mathbf{x}))$. Then $\phi \in \stackrel{\circ}{W}_{2}^{1}(\Omega \cap B)$ and we have

$$
\partial_{x_{j}} w=\sum_{k=1}^{n} \partial_{y_{k}} z \partial_{x_{j}} J_{k}, \quad \partial_{x_{j}} \phi=\sum_{l=1}^{n} \partial_{y_{l}} \psi \partial_{x_{j}} J_{l}
$$

and hence

$$
\int_{\Omega \cap B} \nabla w \nabla \phi d \mathbf{x}=\int_{Q_{+}} \sum_{k, j, l=1}^{n} \partial_{x_{j}} J_{k} \partial_{x_{j}} J_{l} \partial_{y_{k}} z \partial_{y_{l}} \psi\left|\operatorname{det} \mathcal{J}_{H}\right| d \mathbf{y}=\int_{Q_{+}} \sum_{k, l=1}^{n} a_{k, l} \partial_{y_{k}} z \partial_{y_{l}} \psi d \mathbf{y}
$$

where $\mathcal{J}$ is the Jacobi matrix of $H$. We note that we can write

$$
a_{k, l}=\left|\operatorname{det} \mathcal{J}_{H}\right| \mathcal{J}_{J} \mathcal{J}_{J}^{T}
$$

and thus we have

$$
\begin{equation*}
\sum_{k, l=1}^{n} a_{k, l} \xi_{k} \xi_{l}=\left|\operatorname{det} \mathcal{J}_{H}\right|\left(\mathcal{J}_{J}^{T} \boldsymbol{\xi}, \mathcal{J}_{J}^{T} \boldsymbol{\xi}\right) \geq \alpha|\boldsymbol{\xi}|^{2} \tag{1.73}
\end{equation*}
$$

for all $\boldsymbol{\xi}=\left(\xi_{1}, \ldots, \xi_{n}\right) \in \mathbb{R}^{n}$ since both Jacobi matrices $\mathcal{J}_{H}, \mathcal{J}_{J}$ are nonsingular. Also

$$
\int_{\Omega \cap B} g \phi d \mathbf{x}=\int_{Q_{+}}(g \circ H) \psi\left|\operatorname{det} \mathcal{J}_{H}\right| d \mathbf{y}=\int_{Q_{+}} G \psi d \mathbf{y}
$$

where $G \in L_{2}\left(Q_{+}\right)$so that $z \in W_{2}^{1}(Q)$ is a solution to the (elliptic) variational problem

$$
\begin{equation*}
\int_{Q_{+}} \sum_{k, l=1}^{n} a_{k, l} \partial_{y_{k}} z \partial_{y_{l}} \psi d \mathbf{y}=\int_{Q_{+}} G \psi d \mathbf{y}, \quad \psi \in \stackrel{\circ}{W}_{2}^{1}\left(Q_{+}\right) \tag{1.74}
\end{equation*}
$$

Next the process is split into two cases. First we shall consider the method of finite differences, as in the $G_{0}$ case but only in the directions parallel to the boundary. Thus, we take $\psi=D_{-h}\left(D_{h} z\right)$ for $|\mathbf{h}|$ small enough to still have $\psi \in \stackrel{\circ}{W} \frac{1}{2}\left(Q_{+}\right)$. Then, as above

$$
\int_{Q_{+}} D_{h}\left(\sum_{k, l=1}^{n} a_{k, l} \partial_{y_{k}} z\right) \partial_{y_{l}}\left(D_{h} z\right) d \mathbf{y}=\int_{Q_{+}} G D_{-h}\left(D_{h} z\right) d \mathbf{y}
$$

Since $D_{h} x \in \stackrel{o}{W}_{2}^{1}\left(Q_{+}\right)$, we can use Friedrichs lemma to estimate

$$
\int_{Q_{+}} G D_{-h}\left(D_{h} z\right) d \mathbf{y} \leq\|G\|_{0, Q_{+}}\left\|D_{-h}\left(D_{h} z\right)\right\|_{0, Q_{+}} \leq\|G\|_{0, Q_{+}}\left\|\nabla\left(D_{h} z\right)\right\|_{0, Q_{+}}
$$

Then, using $\tau_{h}(f g)-f g=\tau_{h} f\left(\tau_{h} g-g\right)+\left(\tau_{h} f-f\right) g$, we find

$$
D_{h}\left(\sum_{k, l=1}^{n} a_{k, l} \partial_{y_{k}} z\right)(\mathbf{y})=a_{k, l}(\mathbf{y}+\mathbf{h}) \partial_{y_{k}} D_{h} z(\mathbf{y})+\left(D_{h} a_{k, l}\right)(\mathbf{y}) \partial_{y_{k}}(\mathbf{y})
$$

and thus we can write, be the reverse Cauchy-Schwarz inequality

$$
\begin{aligned}
& \int_{Q_{+}} D_{h}\left(\sum_{k, l=1}^{n} a_{k, l} \partial_{y_{k}} z\right) \partial_{y_{l}}\left(D_{h} z\right) d \mathbf{y} \\
& =\int_{Q_{+}} \sum_{k, l=1}^{n}\left(\tau_{h} a_{k, l}\right) \partial_{y_{k}}\left(D_{h} z\right) \partial_{y_{l}}\left(D_{h} z\right) d \mathbf{y}+\int_{Q_{+}} \sum_{k, l=1}^{n}\left(D_{h} a_{k, l}\right) \partial_{y_{k}} z \partial_{y_{l}}\left(D_{h} z\right) d \mathbf{y} \\
& \geq \alpha\left\|\nabla\left(D_{h} z\right)\right\|_{0, Q_{+}}^{2}-C\|\nabla z\|_{0, Q_{+}}\left\|\nabla\left(D_{h} z\right)\right\|_{0, Q_{+}}
\end{aligned}
$$

where $C$ depends on the $C^{1}$ norm of $a_{k, l}$ (and thus $C^{2}$ norm of the local atlas). Thus

$$
\begin{align*}
\left\|\nabla\left(D_{h} z\right)\right\|_{0, Q_{+}}^{2} & \leq \alpha^{-1}\left(\|G\|_{0, Q_{+}}\left\|\nabla\left(D_{h} z\right)\right\|_{0, \Omega}+C\|z\|_{1, \Omega}\left\|\nabla\left(D_{h} z\right)\right\|_{0, Q_{+}}\right) \\
& \leq C^{\prime}\|G\|_{0, Q_{+}}\left\|\nabla\left(D_{h} z\right)\right\|_{0, Q_{+}}, \tag{1.75}
\end{align*}
$$

where we have used the $W_{2}^{1}(\Omega)$ estimates for solutions to (1.74): for $\psi=$ $z \in \stackrel{\mathrm{o}}{ }^{1}{ }_{2}^{1}\left(Q_{+}\right)$

$$
\alpha\|\nabla z\|^{2} \leq \int_{Q_{+}} \sum_{k, l=1}^{n} a_{k, l} \partial_{y_{k}} z \partial_{y_{l}} z d \mathbf{y}=\int_{Q_{+}} G z d \mathbf{y} \leq\|G\|_{0, Q_{+}}\|\nabla z\|_{0, Q_{+}}
$$

Note that in the last inequality we used the Poincarè inequality as $z \in W_{2}^{1}\left(Q_{+}\right)$ and the constant in this inequality can be taken 1 .

Thus we have

$$
\begin{equation*}
\left\|\nabla\left(D_{h} z\right)\right\|_{0, Q_{+}} \leq C^{\prime}\|G\|_{0, Q_{+}}, \tag{1.76}
\end{equation*}
$$

for any $\mathbf{h}$ which is parallel to $Q_{0}$. Let $j=1, \ldots, n, \mathbf{h}=|\mathbf{h}| \mathbf{e}_{k}, k=1, \ldots, n-1$ and $\phi \in C_{0}^{\infty}\left(Q_{+}\right)$. Then we can write

$$
\int_{Q_{+}} D_{h} \partial_{y_{j}} z \phi d \mathbf{y}=-\int_{Q_{+}} \partial_{y_{j}} z D_{-h} \phi d \mathbf{y}
$$

and, by (1.76),

$$
\left|\int_{Q_{+}} \partial_{y_{j}} z D_{-h} \phi d \mathbf{y}\right|=\left|\int_{Q_{+}} D_{h} \partial_{y_{j}} z \phi d \mathbf{y}\right| \leq C^{\prime}\|G\|_{0, Q_{+}}\|\phi\|_{0, Q_{+}}
$$

which, passing to the limit as $|h| \rightarrow 0$ gives for any $(j, k) \neq(n, n)$

$$
\begin{equation*}
\left|\int_{Q_{+}} \partial_{y_{j}} z \partial_{y_{k}} \phi d \mathbf{y}\right| \leq C^{\prime}\|G\|_{0, Q_{+}}\|\phi\|_{0, Q_{+}} . \tag{1.77}
\end{equation*}
$$

To conclude, we have to show also the above estimate for $k=n$. First we observe that $a_{n n} \geq \alpha$ on $Q_{+}$. This follows from (1.73) by taking $\boldsymbol{\xi}=(1,0, \ldots, 0)$. Thus, we can replace in (1.74) $\psi$ by $\psi / a_{n n}$. Then we rewrite (1.74) as

$$
\begin{aligned}
\int_{Q_{+}} a_{n, n} \partial_{y_{k}} z \partial_{y_{l}}\left(a_{n, n}^{-1} \psi\right) d \mathbf{y}= & \int_{Q_{+}} a_{n, n} G\left(a_{n, n}^{-1} \psi\right) d \mathbf{y} \\
& -\int_{Q_{+}} \sum_{(k, l) \neq(n, n)} a_{k, l} \partial_{y_{k}} z \partial_{y_{l}}\left(a_{n, n}^{-1} \psi\right) d \mathbf{y}
\end{aligned}
$$

and differentiating on the left hand side

$$
\begin{aligned}
\int_{Q_{+}} \partial_{y_{k}} z \partial_{y_{l}} \psi d \mathbf{y}= & \int_{Q_{+}} a_{n, n}^{-1} \psi \partial_{y_{n}} a_{n, n} \partial_{y_{k}} z d \mathbf{y}+\int_{Q_{+}} a_{n, n} G \cdot\left(a_{n, n}^{-1} \psi\right) d \mathbf{y} \\
& -\int_{Q_{+}} \sum_{(k, l) \neq(n, n)}\left(a_{n, n}^{-1} \psi\right) \partial_{y_{l}} a_{k, l} \partial_{y_{k}} z d \mathbf{y} \\
& \int_{Q_{+}} \sum_{(k, l) \neq(n, n)} \partial_{y_{k}} z \partial_{y_{l}}\left(a_{n, n}^{-1} a_{k, l} \psi\right) d \mathbf{y}
\end{aligned}
$$

Applying now (1.79), we get

$$
\begin{equation*}
\left|\int_{Q_{+}} \partial_{y_{k}} z \partial_{y_{l}} \psi d \mathbf{y}\right| \leq C\left(\|G\|_{0, Q_{+}}+\|z\|_{1, Q_{+}}\right)\|\psi\|_{0, Q} \tag{1.78}
\end{equation*}
$$

This shows that

$$
\begin{equation*}
\left|\int_{Q_{+}} \partial_{y_{j}} z \partial_{y_{k}} \phi d \mathbf{y}\right| \leq C^{\prime}\|G\|_{0, Q_{+}}\|\phi\|_{0, Q_{+}} . \tag{1.79}
\end{equation*}
$$

for any $j, k=1, \ldots n$ and thus, by Proposition 1.49, each first derivative of $z$ belongs to $W_{2}^{1}\left(Q_{+}\right)$and thus $z \in W_{2}^{2}\left(Q_{+}\right)$. Using the first part of the proof and transferring the solution back to $\Omega$ shows that $u \in W_{2}^{2}(\Omega)$.

Let us consider higher derivatives. As before, we split $u$ according to the partition of unity and separately argue argue in $G_{0} \Subset \Omega$ and in $Q_{+}$. Let us begin with $u \in W_{2}^{2}(\Omega) \cap \stackrel{o}{W}_{2}^{1}(\Omega)$ and consider $w=\beta_{0} u$. Let $f \in W_{2}^{1}(\Omega)$ and consider any derivative $\partial u, i=1, \ldots, n$. We know that $\partial u \in W_{2}^{1}(\Omega)$. Then we can use $\phi \in C_{0}^{\infty}$ and take $\partial \phi$ as the test function in (1.69) so that, integrating by parts

$$
-\int_{\Omega} \partial f \phi d \mathbf{x}=\int_{\Omega} f \partial \phi d \mathbf{x}=\int_{\Omega} \nabla u \nabla \partial \phi d \mathbf{x}=-\int_{\Omega} \nabla \partial u \nabla \phi d \mathbf{x}
$$

so that $\partial u$ is a variational solution with square integrable right hand side and thus $\partial u \in W_{2}^{2}(\Omega)$ and $u \in W_{2}^{3}(\Omega)$. Then we can proceed by induction.

Let us consider $z \in W_{2}^{2}\left(Q_{+}\right) \cap \stackrel{\mathrm{o}}{ }_{W}^{2}$ and let $\partial u$ be any derivative in direction tangential to $Q_{0}$. We claim that $\partial z \in W_{2}^{1}$. First, we note that $D_{h} z \in W_{2}^{\circ}$ if $\mathbf{h}$ is parallel to $Q_{0}$ for sufficiently small $|\mathbf{h}|$. By (1.76), $D_{h} z$ is bounded in $W_{0}^{1}(Q)$ and thus we have a subsequence $\mathbf{h}_{n}$ such that $D_{h_{n}} z \rightharpoonup g \in W_{2}^{1}(Q)$. Clearly, $D_{h_{n}} z$ converges weakly in $L_{2}\left(Q_{+}\right)$and thus for any $\phi \in C_{0}^{\infty}\left(Q_{+}\right)$

$$
\int_{Q_{+}}\left(D_{h_{n}} z\right) \phi d \mathbf{y}=\int_{Q_{+}} z D_{-h_{n}} \phi d \mathbf{y}
$$

and thus passing to the limit

$$
\int_{Q_{+}} g \phi d \mathbf{y}=-\int_{Q_{+}} z \partial \phi d \mathbf{y}
$$

and thus $\partial z \in{ }_{W}^{\circ}{ }_{2}^{1}\left(Q_{+}\right)$. Then, as before

$$
\begin{equation*}
\int_{\Omega} \partial G \psi d \mathbf{y}=\int_{\Omega} \sum_{k, l=1}^{n} \partial_{y_{k}}(\partial z) \partial_{y_{l}} \psi d \mathbf{y} \tag{1.80}
\end{equation*}
$$

for any $\phi \in W_{2}^{1}\left(Q_{+}\right)$. We argue by induction in $m$. Let $f \in W_{2}^{m+1}\left(Q_{+}\right)$. From induction assumption, we have $u \in W^{m+2}\left(Q_{+}\right)$. Also $\partial u$ in any tangential derivative is in ${ }_{W}^{\mathrm{o}} \frac{1}{2}\left(Q_{+}\right)$and satisfies (1.80). By induction assumption to $\partial u$ and $\partial G$ we see that $\partial u \in W_{2}^{m+2}\left(Q_{+}\right)$. Finally we can write

$$
\partial_{x_{n} x_{n}}^{2} u=\frac{1}{a_{n, n}}\left(-G-\int_{\Omega} \sum_{(k, l) \neq(n, n)} \partial_{y_{k}}(\partial z) \partial_{y_{l}} \psi d \mathbf{y}\right)
$$

so that the claim follows.

## An Overview of Semigroup Theory

In this chapter we are concerned with methods of finding solutions of the Cauchy problem.
Definition 2.1. Given a Banach space and a linear operator $\mathcal{A}$ with domain $D(\mathcal{A})$ and range $\operatorname{Im} \mathcal{A}$ contained in $X$ and also given an element $u_{0} \in X$, find a function $u(t)=u\left(t, u_{0}\right)$ such that

1. $u(t)$ is continuous on $[0, \infty)$ and continuously differentiable on $(0, \infty)$, 2. for each $t>0, u(t) \in D(\mathcal{A})$ and

$$
\begin{equation*}
u^{\prime}(t)=\mathcal{A} u(t), \quad t>0 \tag{2.1}
\end{equation*}
$$

3. 

$$
\begin{equation*}
\lim _{t \rightarrow 0^{+}} u(t)=u_{0} \tag{2.2}
\end{equation*}
$$

in the norm of $X$.
A function satisfying all conditions above is called the classical (or strict) solution of (2.1), (2.2).

### 2.1 What the semigroup theory is about

In the theory of differential equations, one of the first differential equations encountered is

$$
\begin{equation*}
u^{\prime}(t)=\alpha u(t), \quad \alpha \in \mathbb{C} \tag{2.3}
\end{equation*}
$$

with initial condition $u(0)=u_{0}$. It is not difficult to verify that $u(t)=e^{t \alpha} u_{0}$ is a solution of Eq. (2.3).

As early as in 1887, G.P. Peano showed that the system of linear ordinary differential equations with constant coefficients

$$
\begin{align*}
u_{1}^{\prime} & =\alpha_{11} u_{1}+\cdots+\alpha_{1 n} u_{n} \\
& \vdots  \tag{2.4}\\
u_{n}^{\prime} & =\alpha_{n 1} u_{1}+\cdots+\alpha_{n n} u_{n}
\end{align*}
$$

can be written in a matrix form as

$$
\begin{equation*}
u^{\prime}(t)=A u(t) \tag{2.5}
\end{equation*}
$$

where $A$ is an $n \times n$ matrix $\left\{\alpha_{i j}\right\}_{1 \leq i, j \leq n}$ and $u$ is an $n$-vector whose components are unknown functions, and can be solved using the explicit formula

$$
\begin{equation*}
u(t)=e^{t A} u_{0} \tag{2.6}
\end{equation*}
$$

where the matrix exponential $e^{t A}$ is defined by

$$
\begin{equation*}
e^{t A}=I+\frac{t A}{1!}+\frac{t^{2} A^{2}}{2!}+\cdots \tag{2.7}
\end{equation*}
$$

Taking a norm on $\mathbb{C}^{n}$ and the corresponding matrix-norm on $M_{n}(\mathbb{C})$, the space of all complex $n \times n$ matrices, one shows that the partial sums of the series (2.7) form a Cauchy sequence and converge. Moreover, the map $t \rightarrow e^{t A}$ is continuous and satisfies the properties, [79, Proposition I.2.3]:

$$
\begin{array}{ll}
e^{(t+s) A} & =e^{t A} e^{s A} \quad \text { for all } t, s \geq 0 \\
e^{0 A} & =I \tag{2.8}
\end{array}
$$

Thus the one-parameter family $\left\{e^{t A}\right\}_{t \geq 0}$ is a homomorphism of the additive semigroup $[0, \infty)$ into a multiplicative semigroup of matrices $M_{n}$ and forms what is termed a semigroup of matrices.

The representation (2.7) can be used to obtain a solution of the abstract Cauchy problem (2.1-2.2) where $\mathcal{A}: X \rightarrow X$ is a bounded linear operator, as in this case the series in (2.7) is still convergent with respect to the norm in the space of linear operators $\mathcal{L}(X)$.

In general, however, the operators coming from applications, such as, for example, differential operators, are not bounded on the whole space $X$ and (2.7) cannot be used to obtain a solution of the abstract Cauchy problem (??). This is due to the fact that the domain of the operator $\mathcal{A}$ in such cases is a proper subspace of $X$ and because (2.7) involves iterates of $A$, their common domain could shrink to the trivial subspace $\{0\}$. For the same reason, another common representation of the exponential function

$$
\begin{equation*}
e^{t \mathcal{A}}=\lim _{n \rightarrow \infty}\left(1+\frac{t}{n} \mathcal{A}\right)^{n} \tag{2.9}
\end{equation*}
$$

cannot be used. For a large class of unbounded operators a variation of the latter, however, makes the representation (2.6) meaningful with $e^{t \mathcal{A}}$ calculated according to the formula

$$
\begin{equation*}
e^{t \mathcal{A}} x=\lim _{n \rightarrow \infty}\left(I-\frac{t}{n} \mathcal{A}\right)^{-n} x=\lim _{n \rightarrow \infty}\left[\frac{n}{t}\left(\frac{n}{t}-\mathcal{A}\right)^{-1}\right]^{n} x \tag{2.10}
\end{equation*}
$$

The aim of the semigroup theory is to find conditions under which such a generalization of the exponential function satisfying (2.8) is possible.

### 2.2 Rudiments

### 2.2.1 Definitions and Basic Properties

If the solution to $(2.1),(2.2)$ is unique, then we can introduce the family of operators $(G(t))_{t \geq 0}$ such that $u\left(t, u_{0}\right)=G(t) u_{0}$. Ideally, $G(t)$ should be defined on the whole space for each $t>0$, and the function $t \rightarrow G(t) u_{0}$ should be continuous for each $u_{0} \in X$, leading to well-posedness of (2.1), (2.2). Moreover, uniqueness and linearity of $\mathcal{A}$ imply that $G(t)$ are linear operators. A fine-tuning of these requirements leads to the following definition.

Definition 2.2. A family $(G(t))_{t \geq 0}$ of bounded linear operators on $X$ is called a $C_{0}$-semigroup, or a strongly continuous semigroup, if
(i) $G(0)=I$;
(ii) $G(t+s)=G(t) G(s)$ for all $t, s \geq 0$;
(iii) $\lim _{t \rightarrow 0^{+}} G(t) x=x$ for any $x \in X$.

A linear operator $A$ is called the (infinitesimal) generator of $(G(t))_{t \geq 0}$ if

$$
\begin{equation*}
A x=\lim _{h \rightarrow 0^{+}} \frac{G(h) x-x}{h} \tag{2.11}
\end{equation*}
$$

with $D(A)$ defined as the set of all $x \in X$ for which this limit exists. If we need to use differen generators, then typically the semigroup generated by $A$ will be denoted by $\left(G_{A}(t)\right)_{t \geq 0}$, otherwise simply by $(G(t))_{t \geq 0}$.

## Why $C_{0}$-semigroups?

Proposition 2.3. If $(G(t))_{t \geq 0}$ is uniformly bounded, then its generator is bounded.

Proof. Since $\rho^{-1} \int_{0}^{\rho} G(s) d s \rightarrow I$ in the uniform operator norm, then there is $\rho>0$ such that $\left\|\rho^{-1} \int_{0}^{\rho} G(s) d s-I\right\|<1$ and thus $\rho^{-1} \int_{0}^{\rho} G(s) d s$ and hence $\int_{0}^{\rho} G(s) d s$ are invertible.

$$
\begin{aligned}
\frac{G(h)-I}{h} \int_{0}^{\rho} G(s) d s & =\frac{1}{h} \int_{0}^{\rho}(G(s+h)-G(s)) d s \\
& =\frac{1}{h} \int_{\rho}^{\rho+h} G(s) x d s-\frac{1}{h} \int_{0}^{\rho} G(s) x d s
\end{aligned}
$$

Thus

$$
\frac{G(h)-I}{h}=\left(\frac{1}{h} \int_{\rho}^{\rho+h} G(s) x d s-\frac{1}{h} \int_{0}^{\rho} G(s) x d s\right)\left(\int_{0}^{\rho} G(s) d s\right)^{-1}
$$

Letting $h \rightarrow 0$, we see that $(G(h)-I) / h \rightarrow(G(\rho)-I)\left(\int_{0}^{\rho} G(s) d s\right)^{-1}$ in the uniform norm and thus the generator is bounded.

Proposition 2.4. Let $(G(t))_{t \geq 0}$ be a $C_{0}$-semigroup. Then there are constants $\omega \geq 0, M \geq 1$ such that

$$
\begin{equation*}
\|G(t)\| \leq M e^{\omega t}, \quad t \geq 0 \tag{2.12}
\end{equation*}
$$

Proof. First we observe that $\|G(t)\|$ is bounded on some interval. Indeed, if not, there is $\left(t_{n}\right)_{n \in \mathbb{N}}, t_{n} \rightarrow 0,\left\|G\left(t_{n}\right)\right\| \geq n$, that is $\left(G\left(t_{n}\right)\right)$ is unbounded. But, by the Banach-Steinhaus theorem there is an $x \in X$ and a subsequence $\left(t_{n_{k}}\right)_{n_{k} \in \mathbb{N}}$ such that $\left(G\left(t_{n_{k}}\right) x\right)$ is unbounded, contrary to (iii). So, $\|G(t)\| \leq M$ for $0 \leq t \leq \eta$ for some $\eta$ and $M \geq 1$ as $G(0)=I$. For any $t \geq 0$ we take $t=n \eta+\delta, 0 \leq \delta<\eta$ and, by the semigroup property,

$$
\|G(t)\|=\left\|G(\delta)(G(\eta))^{n}\right\| \leq M M^{n}=M e^{(t-\delta) \ln M / \eta} \leq M e^{\omega t}
$$

where $\omega=\eta^{-1} \ln M \geq 0$.
As a corollary, we have
Corollary 2.5. Let $(G(t))_{t \geq 0}$ be a $C_{0}$-semigroup. Then for every $x \in X$, $t \rightarrow G(t) x \in C\left(\mathbb{R}_{+} \cup\{0\}, X\right)$.

Proof. We have for $t, h \geq 0$

$$
\|G(t+h) x-G(t) x\| \leq\|G(t)\|\|G(h) x-x\| \leq M e^{\omega t}\|G(h) x-x\|
$$

and for $t \geq h \geq 0$

$$
\|G(t-h) x-G(t) x\| \leq\|G(t-h)\|\|G(h) x-x\| \leq M e^{\omega t}\|G(h) x-x\|
$$

and the statement follows from condition (iii).
Remark 2.6. As we have seen above, for semigroups, the existence of a onesided limit at some $t_{0}>0$ yields the existence of the limit.

Let $(G(t))_{t \geq 0}$ be a semigroup generated by the operator $A$. The following properties of $(G(t))_{t \geq 0}$ are frequently used.

Lemma 2.7. Let $(G(t))_{t \geq 0}$ be a $C_{0}$-semigroup generated by $A$.
(a) For $x \in X$

$$
\begin{equation*}
\lim _{h \rightarrow 0} \frac{1}{h} \int_{t}^{t+h} G(s) x d s=G(t) x \tag{2.13}
\end{equation*}
$$

(b) For $x \in X, \int_{0}^{t} G(s) x d s \in D(A)$ and

$$
\begin{equation*}
A \int_{0}^{t} G(s) x d s=G(t) x-x \tag{2.14}
\end{equation*}
$$

(c) For $x \in D(A), G(t) x \in D(A)$ and

$$
\begin{equation*}
\frac{d}{d t} G(t) x=A G(t) x=G(t) A x \tag{2.15}
\end{equation*}
$$

(d) For $x \in D(A)$,

$$
\begin{equation*}
G(t) x-G(s) x=\int_{s}^{t} G(\tau) A x d \tau=\int_{s}^{t} A G(\tau) x d \tau \tag{2.16}
\end{equation*}
$$

Proof. (a) follows from continuity of the semigroup. To prove (b) we consider $x \in X$ and $h>0$. Then

$$
\begin{aligned}
\frac{G(h)-I}{h} \int_{0}^{t} G(s) x d s & =\frac{1}{h} \int_{0}^{t}(G(s+h) x-G(s) x) d s \\
& \left.=\frac{1}{h} \int_{t}^{t+h} G(s) x d s-\frac{1}{h} \int_{0}^{t} G(s) x\right) d s
\end{aligned}
$$

and the right hand side tends to $G(t) x-x$ by (a) which proves that $\int_{0}^{t} G(s) x d s \in D(A)$ and (2.14). To prove (c), let $x \in D(A)$ and $h>0$. As above

$$
\frac{G(h)-I}{h} G(t) x=G(t)\left(\frac{G(h)-I}{h}\right) x \rightarrow T(t) x
$$

as $h \rightarrow 0$. Thus, $G(t) x \in D(A)$ and $A G(t) x=G(t) A x$ for $x \in D(A)$. The limit above also shows that

$$
\frac{d^{+}}{d t} G(t) x=A G(t) x=G(t) A x
$$

that is, the right derivative of $G(t) x$ is $A G(t)$. Take now $t>0$ and $h \leq t$. Then

$$
\begin{aligned}
& \lim _{h \rightarrow 0}\left(\frac{G(t-h) x-G(t) x}{-h}-A G(t) x\right) \\
& \lim _{h \rightarrow 0} G(t-h)\left(\frac{G(h) x-x}{h}-A x\right)+\lim _{h \rightarrow 0}(G(t-h) A x-G(t) A x)
\end{aligned}
$$

and we see that both limits are 0 by uniform boundedness of $(G(t))_{t \geq 0}$, strong continuity and $x \in D(A)$.

Part (d) is obtained by integrating (2.15).

From (2.15) and condition (iii) of Definition 2.2 we see that if $A$ is the generator of $(G(t))_{t \geq 0}$, then for $x \in D(A)$ the function $t \rightarrow G(t) x$ is a classical solution of the following Cauchy problem,

$$
\begin{align*}
\partial_{t} u(t) & =A(u(t)), \quad t>0  \tag{2.17}\\
\lim _{t \rightarrow 0^{+}} u(t) & =x \tag{2.18}
\end{align*}
$$

We note that ideally the generator $A$ should coincide with $\mathcal{A}$ but in reality very often it is not so.

Remark 2.8. We noted above that for $x \in D(A)$ the function $u(t)=G(t) x$ is a classical solution to (2.17), (2.18). For $x \in X \backslash D(A)$, however, the function $u(t)=G(t) x$ is continuous but, in general, not differentiable, nor $D(A)$-valued, and, therefore, not a classical solution. Nevertheless, from (2.14), it follows that the integral $v(t)=\int_{0}^{t} u(s) d s \in D(A)$ and therefore it is a strict solution of the integrated version of (2.17), (2.18):

$$
\begin{align*}
\partial_{t} v & =A v+x, \quad t>0 \\
v(0) & =0 \tag{2.19}
\end{align*}
$$

or equivalently,

$$
\begin{equation*}
u(t)=A \int_{0}^{t} u(s) d s+x \tag{2.20}
\end{equation*}
$$

We say that a function $u$ satisfying (2.19) (or, equivalently, (2.20)) is a mild solution or integral solution of (2.17), (2.18).

Corollary 2.9. If $(G(t))_{t \geq 0}$ is a $C_{0}$-semigroup generated by $A$, then $A$ is a closed densely defined linear operator.

Proof. For $x \in X$ we set $x_{t}=t^{-1} \int_{0}^{t} G(s) x d s$. By (b), $x_{t} \in D(A)$ and by (a), $x_{t} \rightarrow x$ as $t \rightarrow 0$. To prove closedness, let $D(A) \ni x_{n} \rightarrow x \in X$ and let $A x_{n} \rightarrow y \in X$. From (d) we have

$$
G(t) x_{n}-x_{n}=\int_{0}^{t} G(s) A x_{n} d s
$$

By local boundedness of $(G(t))_{t \geq 0}$ we have that $G(s) A x_{n} \rightarrow G(s) y$ uniformly on bounded intervals, hence, by letting $n \rightarrow \infty$,

$$
G(t) x-x=\int_{0}^{t} G(s) y d s
$$

Thus, using (a), $x \in D(A)$ and $A x \in y$.

Thus, if we have a semigroup, we can identify the Cauchy problem of which it is a solution. Usually, however, we are interested in the reverse question, that is, in finding the semigroup for a given equation.

A first step in this direction is
Theorem 2.10. Let $\left(G_{A}(t)\right)_{t>0}$ and $\left(G_{B}(t)\right)_{t>0}$ be $C_{0}$ semigroups generated by, respectively, $A$ and $B$. If $\bar{A}=B$, then $G_{A}(\bar{t})=G_{B}(T)$.

Proof. Let $x \in D(A)=D(B)$. Consider the function

$$
s \rightarrow G_{A}(t-s) G_{B}(s) x, \quad 0 \leq s \leq t
$$

is continuous on $[0, t]$. Writing, for appropriate $s, h$

$$
\begin{aligned}
& \frac{G_{A}(t-(s+h)) G_{B}(s+h) x-G_{A}(t-s) G_{B}(s) x}{h} \\
& =\frac{G_{A}(t-(s+h)) G_{B}(s+h) x-G_{A}(t-(s+h)) G_{B}(s) x}{h} \\
& \quad+\frac{G_{A}(t-(s+h)) G_{B}(s) x-G_{A}(t-s) G_{B}(s) x}{h}
\end{aligned}
$$

we see that by local boundedness both terms converge and, by (c), we obtain

$$
\begin{aligned}
\frac{d}{d s} G_{A}(t-s) G_{B}(s) x & =-A G_{A}(t-s) G_{B}(s) x+G_{A}(t-s) B G_{B}(s) x \\
& =-G_{A}(t-s) A G_{B}(s) x+G_{A}(t-s) B G_{B}(s) x=0
\end{aligned}
$$

Thus $G_{A}(t-s) G_{B}(s) x$ is constant and, in particular, evaluating at $s=0$ and $s=t$ we get $G_{A}(t) x=G_{B}(t) x$ for any $t$ and $x \in D(A)$. From density, we obtain the equality on $X$.

The final answer is given by the Hille-Yoshida theorem (or, more properly, the Feller-Miyadera-Hille-Phillips-Yosida theorem). Before, however, we need to discuss the concept of resolvent.

Let $A$ be any operator in $X$. The resolvent set of $A$ is defined as

$$
\begin{equation*}
\rho(A)=\{\lambda \in \mathbb{C} ; \lambda I-A: D(A) \rightarrow X \text { is invertible }\} \tag{2.21}
\end{equation*}
$$

We call $(\lambda I-A)^{-1}$ the resolvent of $A$ and denote it by $R(\lambda, A)=(\lambda I-A)^{-1}$, $\lambda \in \rho(A)$. The complement of $\rho(A)$ in $\mathbb{C}$ is called the spectrum of $A$ and denoted by $\sigma(A)$.

The resolvent of any operator $A$ satisfies the resolvent identity

$$
\begin{equation*}
R(\lambda, A)-R(\mu, A)=(\mu-\lambda) R(\lambda, A) R(\mu, A), \quad \lambda, \mu \in \rho(A) \tag{2.22}
\end{equation*}
$$

from which it follows, in particular, that $R(\lambda, A)$ and $R(\mu, A)$ commute. Writing

$$
R(\mu, A)=R(\lambda, A)(I-(\mu-\lambda) R(\mu, A))
$$

we see by the Neuman expansion that $R(\lambda, A)$ can be written as the power series

$$
\begin{equation*}
R(\lambda, A)=\sum_{n=0}^{\infty}(\mu-\lambda)^{n} R(\mu, A)^{n+1} \tag{2.23}
\end{equation*}
$$

for $|\mu-\lambda|<\|R(\mu, A)\|^{-1}$ so that $\rho(A)$ is open and $\lambda \rightarrow R(\lambda, A)$ is an analytic function in $\rho(A)$. The iterates of the resolvent and its derivatives are related by

$$
\begin{equation*}
\frac{d^{n}}{d \lambda^{n}} R(\lambda, A)=(-1)^{n} n!R(\lambda, A)^{n+1} \tag{2.24}
\end{equation*}
$$

### 2.2.2 The Hille-Yosida Theorem

We begin with the simplest case of contractive semigroups. A $C_{0}$ semigroup $\left(G_{A}(t)\right)_{t \geq 0}$ is called contractive if

$$
\left\|G_{A}(T)\right\| \leq 1
$$

Theorem 2.11. $A$ is the generator of a contractive semigroup $\left(G_{A}(t)\right)_{t \geq 0}$ if and only if
(a) $A$ is closed and densely defined,
(b) $(0, \infty) \subset \rho(A)$ and for all $\lambda>0$,

$$
\begin{equation*}
\|R(\lambda, A)\| \leq \frac{1}{\lambda} \tag{2.25}
\end{equation*}
$$

Proof. (Necessity) If $A$ is the generator of a $C_{0}$ semigroup $\left(G_{A}(t)\right)_{t \geq 0}$, then it is densely defined and closed. Let us define

$$
\begin{equation*}
R(\lambda) x=\int_{0}^{\infty} e^{-\lambda t} G(t) x d t \tag{2.26}
\end{equation*}
$$

is valid for all $x \in X$. Since $\left(G_{A}(t)\right)_{t \geq 0}$ is contractive, the integral exists for $\lambda>0$ as an improper Riemann integral and defines a bounded linear operator $R(\lambda) x$ (by the Banach-Steinhaus theorem). $R(\lambda)$ satisfeis

$$
\|R(\lambda) x\| \leq \frac{1}{\lambda}\|x\|
$$

Furthermore, $h>0$,

$$
\begin{aligned}
\frac{G_{A}(t)-I}{h} R(\lambda) x & =\frac{1}{h} \int_{0}^{\infty} e^{-\lambda t}\left(G_{A}(t+h) x-G_{A}(t) x\right) d t \\
& =\frac{1}{h}\left(\int_{h}^{\infty} e^{-\lambda(t-h)} G_{A}(t) x d t-\int_{0}^{\infty} e^{-\lambda t} G_{A}(t) x d t\right) \\
& =\frac{e^{\lambda h}-1}{h} \int_{h}^{\infty} e^{-\lambda t} G_{A}(t) x d t-\frac{e^{\lambda h}}{h} \int_{0}^{\infty} e^{-\lambda t} G_{A}(t) x d t
\end{aligned}
$$

By strong continuity of $G_{A}$, the right hand side converges to $\lambda R(\lambda) x-x$. This implies that for any $x \in D(A)$ and $\lambda>0$ we have $R(\lambda) x \in D(A)$ and $A R(\lambda)=\lambda R(\lambda)-I$ so

$$
\begin{equation*}
(\lambda I-A) R(\lambda)=I \tag{2.27}
\end{equation*}
$$

On the other hand, for $x \in D(A)$ we have

$$
R(\lambda) A x=\int_{0}^{\infty} e^{-\lambda t} G(t) A x d t=A\left(\int_{0}^{\infty} e^{-\lambda t} G(t) x\right) d t=A R(\lambda) x
$$

by commutativity (Lemma 2.7 (c)) and closedness of $A$. Thus $A$ and $R(\lambda)$ commute and

$$
R(\lambda)(\lambda I-A) x=A x
$$

on $D(A)$. Thus $R(\lambda)$ is the resolvent of $A$ and satisfies the desired estimate
The converse is more difficult to prove. The starting point of the second part of the proof is the observation that if $(A, D(A))$ is a closed and densely defined operator satisfying $\rho(A) \supset(0, \infty)$ and $\|\lambda R(\lambda, A)\| \leq 1$ for all $\lambda>0$, then
(i) for any $x \in X$,

$$
\begin{equation*}
\lim _{\lambda \rightarrow \infty} \lambda R(\lambda, A) x=x \tag{2.28}
\end{equation*}
$$

Indeed, first consider $x \in D(A)$. Then

$$
\|\lambda R(\lambda, A) x-x\|=\|A R(\lambda, A) x\|=\|R(\lambda, A) A x\| \leq \frac{1}{\lambda}\|A x\| \rightarrow 0
$$

as $\lambda \rightarrow \infty$. Since $D(A)$ is dense and $\|\lambda R(\lambda, A)\| \leq 1$ then by $3 \epsilon$ argument we extend the convergence to $X$.
(ii) $A R(\lambda, A)$ are bounded operators and for any $x \in D(A)$,

$$
\begin{equation*}
\lim _{\lambda \rightarrow \infty} \lambda A R(\lambda, A) x=A x \tag{2.29}
\end{equation*}
$$

Boundedness follows from $A R(\lambda, A)=\lambda R(\lambda, A)-I$. Eq. (2.29) follows (2.28).

It was Yosida's idea to use the bounded operators

$$
\begin{equation*}
A_{\lambda}=\lambda A R(\lambda, A) \tag{2.30}
\end{equation*}
$$

as an approximation of $A$ for which we can define semigroups uniformly continuous semigroups $\left(G_{\lambda}(t)\right)_{t \geq 0}$ via the exponential series. First we note that $\left(G_{\lambda}(t)\right)_{t \geq 0}$ are semigroups of contractions and, for any $x \in X$ and $\lambda, \mu>0$ we have

$$
\begin{equation*}
\left\|G_{\lambda}(t) x-G_{\mu}(t) x\right\| \leq t\left\|A_{\lambda} x-A_{\mu} x\right\| \tag{2.31}
\end{equation*}
$$

Indeed, using $A_{\lambda}=\lambda^{2} R(\lambda, A)-\lambda I$ and the series estimates

$$
\left\|G_{\lambda} x\right\| \leq e^{-\lambda t} e^{\lambda\|R(\lambda, A)\| t} \leq 1
$$

Further, from the definition operators $G_{\lambda}(t), G_{\mu}(t), A_{\lambda}, A_{\mu}$ commute with each other. Then

$$
\begin{aligned}
\left\|G_{\lambda}(t) x-G_{\mu}(t) x\right\| & =\left\|\int_{0}^{1} \frac{d}{d s} e^{t s A_{\lambda}} e^{t(1-s) A_{\mu}} x d s\right\| \\
& \leq t \int_{0}^{1}\left\|e^{t s A_{\lambda}} e^{t(1-s) A_{\mu}}\left(A_{\lambda} x-A_{\mu} x\right)\right\| d s \leq t\left\|A_{\lambda} x-A_{\mu} x\right\| .
\end{aligned}
$$

Using (2.31) we obtain for $x \in D(A)$

$$
\left\|G_{\lambda}(t) x-G_{\mu}(t) x\right\| \leq t\left\|A_{\lambda} x-A_{\mu} x\right\| \leq t\left(\left\|A_{\lambda} x-A x\right\|+\left\|A x-A_{\mu} x\right\|\right) .
$$

Hence $\left.\left(G_{\lambda}(t) x\right)\right)_{\lambda}$ strongly converges and the convergence (for each $x$ ) is uniform in $t$ on bounded intervals (almost uniform on $\overline{\mathbb{R}_{+}}$. Since $D(A)$ is dense in $X$ and $\left\|G_{\lambda}(t)\right\| \leq 1$ we get

$$
\lim _{\lambda \rightarrow \infty} G_{\lambda}(t) x=: S(t) x
$$

for $x \in X$. The convergence is still almost uniform on $\overline{\mathbb{R}_{+}}$. From the limit we see that $(S(t))_{t \geq 0}$ is a $C_{0}$ semigroup of contractions.

What remains is to show that $(S(t))_{t \geq 0}$ is generated by $A$. Let $x \in D(A)$. Then

$$
\begin{equation*}
G(t) x-x=\lim _{\lambda \rightarrow \infty}\left(G_{\lambda}(t) x-x\right)=\lim _{\lambda \rightarrow \infty} \int_{0}^{t} e^{s A_{\lambda}} A_{\lambda} x d s=\int_{0}^{t} G(s) A x d s \tag{2.32}
\end{equation*}
$$

where the last equality follows from

$$
\begin{aligned}
\left\|e^{s A_{\lambda}} A_{\lambda} x-G(s) A x\right\| & \leq\left\|e^{s A_{\lambda}} A_{\lambda} x-e^{s A_{\lambda}} A x\right\|+\left\|e^{s A_{\lambda}} A x-G(s) A x\right\| \\
& \leq\left\|A_{\lambda} x-A x\right\|+\left\|e^{s A_{\lambda}} A x-G(s) A x\right\|
\end{aligned}
$$

by contractivity of $\left(G_{\lambda}(t)\right)_{t \geq 0}$, so that convergence is uniform on bounded intervals. Assume now that $(G(t))_{t \geq 0}$ is generated by $B$. Dividing (2.32) by $t$ and passing to the limit, we obtain

$$
B x=A x, \quad x \in D(A)
$$

so that $A \subset B$. On the other hand, we know that $I-A$ and $I-B$ are bijections from, resp $D(A)$ and $D(B)$ with $D(A) \subset D(B)$. But then we have $(I-B) D(A)=(I-A) D(A)=X$, that is, $D(A)=(I-B)^{-1} X=D(B)$ so $A=B$.

Corollary 2.12. A linear operator $A$ is the generator of a $C_{0}$ semigroup $(G(t))_{t \geq 0}$ satisfying $\|G(t)\| \leq e^{\omega t}$ if and only if
(i) $A$ is closed and $\overline{D(A)}=X$;
(ii) $\rho(A) \supset(\omega, \infty)$ and for such $\lambda$

$$
\begin{equation*}
\|R(\lambda, A)\| \leq \frac{1}{\lambda-\omega} \tag{2.33}
\end{equation*}
$$

Proof. Follows from the contractive semigroup $S(t)=e^{-\omega t} G(t)$ being generated by $A-\omega I$.

The full version of the Hille-Yosida theorem reads
Theorem 2.13. $A \in \mathcal{G}(M, \omega)$ if and only if
(a) A is closed and densely defined,
(b) there exist $M>0, \omega \in \mathbb{R}$ such that $(\omega, \infty) \subset \rho(A)$ and for all $n \geq 1, \lambda>\omega$,

$$
\begin{equation*}
\left\|(\lambda I-A)^{-n}\right\| \leq \frac{M}{(\lambda-\omega)^{n}} \tag{2.34}
\end{equation*}
$$

### 2.2.3 Dissipative operators and the Lumer-Phillips theorem

Let $X$ be a Banach space (real or complex) and $X^{*}$ be its dual. From the Hahn-Banach theorem, Theorem 1.12 for every $x \in X$ there exists $x^{*} \in X^{*}$ satisfying

$$
<x^{*}, x>=\|x\|^{2}=\left\|x^{*}\right\|^{2} .
$$

Therefore the duality set

$$
\begin{equation*}
\left.\mathcal{J}(x)=\left\{x^{*} \in X^{*} ;<x^{*}, x\right\rangle=\|x\|^{2}=\left\|x^{*}\right\|^{2}\right\} \tag{2.35}
\end{equation*}
$$

is nonempty for every $x \in X$.
Definition 2.14. We say that an operator $(A, D(A))$ is dissipative if for every $x \in D(A)$ there is $x^{*} \in \mathcal{J}(x)$ such that

$$
\begin{equation*}
\Re<x^{*}, A x>\leq 0 \tag{2.36}
\end{equation*}
$$

If $X$ is a real space, then the real part in the above definition can be dropped.

Theorem 2.15. A linear operator $A$ is dissipative if and only if for all $\lambda>0$ and $x \in D(A)$,

$$
\begin{equation*}
\|(\lambda I-A) x\| \geq \lambda\|x\| \tag{2.37}
\end{equation*}
$$

Proof. Let $A$ be dissipative, $\lambda>0$ and $x \in D(A)$. If $x^{*} \in \mathcal{J}$ and $\Re<$ $A x, x^{*}>\leq 0$, then

$$
\|\lambda x-A x\|\|x\| \geq \mid\left(\lambda x-A x, x^{*}>\mid \geq \Re<\lambda x-A x, x^{*}>\geq \lambda\|x\|^{2}\right.
$$

so that we get (2.65).
Conversely, let $x \in D(A)$ and $\lambda\|x\| \leq\|\lambda x-A x\|$ for $\lambda>0$. Consider $y_{\lambda}^{*} \in \mathcal{J}(\lambda x-A x)$ and $z_{\lambda}^{*}=y_{\lambda}^{*} /\left\|y_{\lambda}^{*}\right\|$.

$$
\begin{aligned}
\lambda\|x\| \leq\|\lambda x-A x\| & =\|\lambda x-A x\|\left\|z_{\lambda}^{*}\right\|=\left\|y_{\lambda}^{*}\right\|^{1}\|\lambda x-A x\|\left\|y_{\lambda}^{*}\right\|=\left\|y_{\lambda}^{*}\right\|^{1}<\lambda x-A x, y_{\lambda}^{*}> \\
& =<\lambda x-A x, z_{\lambda}^{*}>=\lambda \Re<x, z_{\lambda}^{*}>-\Re<A x, z_{\lambda}^{*}> \\
& \leq \lambda\|x\|-\Re<A x, z_{\lambda}^{*}>
\end{aligned}
$$

for every $\lambda>0$. From this estimate we obtain that $\Re<A x, z_{\lambda}^{*}>\leq 0$ and, by $|\alpha| \geq \Re \alpha$,
$\lambda \Re<x, z_{\lambda}^{*}>=\lambda\|x\|+\Re<A x, z_{\lambda}^{*}>\geq \lambda\|x\|-\left|\Re<A x, z_{\lambda}^{*}>\right| \geq \lambda\|x\|-\|A x\|$
or $\Re<x, z_{\lambda}^{*}>\geq\|x\|-\lambda^{-1}\|A x\|$. Now, the unit ball in $X^{*}$ is weakly-* compact and thus there is a sequence $\left(z_{\lambda_{n}}^{*}\right)_{n \in \mathbb{N}}$ converging to $z^{*}$ with $\left\|z^{*}\right\|=1$. From the above estimates, we get

$$
\Re<A x, z^{*}>\leq 0
$$

and $\Re<x, z^{*}>\geq\|x\|$. Hence, also, $\left|<x, z^{*}>\right| \geq\|x\|$ On the other hand, $R e<x, z^{*}>\leq\left|<x, z^{*}>\right| \leq\|x\|$ and hence $<x, z^{*}>=\|x\|$. Taking $x^{*}=z^{*}\|x\|$ we see that $x^{*} \in \mathcal{J}(x)$ and $\Re<A x, x^{*}>\leq 0$ and thus $A$ is dissipative.

Theorem 2.16. Let $A$ be a linear operator with dense domain $D(A)$ in $X$.
(a) If $A$ is dissipative and there is $\lambda_{0}>0$ such that the range $\operatorname{Im}\left(\lambda_{0} I-A\right)=$ $X$, then $A$ is the generator of a $C_{0}$-semigroup of contractions in $X$.
(b) If $A$ is the generator of a $C_{0}$ semigroup of contractions on $X$, then $\operatorname{Im}(\lambda I-A)=X$ for all $\lambda>0$ and $A$ is dissipative. Moreover, for every $x \in D(A)$ and every $x^{*} \in \mathcal{J}(x)$ we have $\Re<A x, x^{*}>\leq 0$.

Proof. Let $\lambda>0$, then dissipativeness of $A$ implies $\|\lambda x-A x\| \geq \lambda\|x\|$ for $x \in D(A), \lambda>0$. This gives injectivity and, since by assumption, the $\operatorname{Im}\left(\lambda_{0} I-\right.$ A) $D(A)=X,\left(\lambda_{0} I-A\right)^{-1}$ is a bounded everywhere defined operator and thus closed. But then $\lambda_{0} I-A$, and hence $A$, are closed. We have to prove that $\operatorname{Im}(\lambda I-A) D(A)=X$ for all $\lambda>0$. Consider the set $\Lambda=\{\lambda>0 ; \operatorname{Im}(\lambda I-$ A) $D(A)=X\}$. Let $\lambda \in \Lambda$. This means that $\lambda \in \rho(A)$ and, since $\rho(A)$ is open, $\Lambda$ is open in the induced topology. We have to prove that $\Lambda$ is closed in the induced topology. Assume $\lambda_{n} \rightarrow \lambda, \lambda>0$. For every $y \in X$ there is $x_{n} \in D(A)$ such that

$$
\lambda_{n} x_{n}-A x_{n}=y
$$

From (??), $\left\|x_{n}\right\| \leq \frac{1}{\lambda_{n}}\|y\| \leq C$ for some $C>0$. Now

$$
\begin{aligned}
\lambda_{m}\left\|x_{n}-x_{m}\right\| & \leq\left\|\lambda_{m}\left(x_{n}-x_{m}\right)-A\left(x_{n}-x_{m}\right)\right\| \\
& =\left\|-\lambda_{m} x_{n}+\lambda_{m} x_{m}-\lambda_{n} x_{n}+\lambda_{n} x_{n}-A x_{n}+A x_{m}\right\| \\
& =\left|\lambda_{n}-\lambda_{m}\right|\left\|x_{n}\right\| \leq C\left|\lambda_{n}-\lambda_{m}\right|
\end{aligned}
$$

Thus, $\left(x_{n}\right)_{n \in \mathbb{N}}$ is a Cauchy sequence. Let $x_{n} \rightarrow x$, then $A x_{n} \rightarrow \lambda x-y$. Since $A$ is closed, $x \in D(A)$ and $\lambda x-A x=y$. Thus, for this $\lambda, \operatorname{Im}(\lambda I-A) D(A)=X$ and $\lambda \in \Lambda$. Thus $\Lambda$ is also closed in $(0, \infty)$ and since $\lambda_{0} \in \Lambda, \Lambda \neq \emptyset$ and thus $\Lambda=(0, \infty)$ (as the latter is connected). Thus, the thesis follows from the Hille-Yosida theorem.

On the other hand, if $A$ is the generator of a semigroup of contractions $(G(t))_{t \geq 0}$, then $(0, \infty) \subset \rho(A)$ and $\operatorname{Im}(\lambda I-A) D(A)=X$ for all $\lambda>0$. Furthermore, if $x \in D(A), x^{*} \in \mathcal{J}(x)$, then

$$
\left|<G(t) x, x^{*}>\right| \leq\|G(t) x\|\left\|x^{*}\right\| \leq\|x\|^{2}
$$

and therefore

$$
\Re<G(t) x-x, x^{*}>=\Re<G(t) x, x^{*}>-\|x\|^{2} \leq 0
$$

and, dividing the left hand side by $t$ and passing with $t \rightarrow \infty$, we obtain

$$
<A x, x^{*}>\leq 0 .
$$

Since this holds for every $x^{*} \in \mathcal{J}(x)$, the proof is complete.
Adjoint operators
Before we move to an important corollary, let as recall the concept of the adjoint operator. If $A \in \mathcal{L}(X, Y)$, then the adjoint operator $A^{*}$ is defined as

$$
\begin{equation*}
<y^{*}, A x>=<A^{*} y^{*}, x> \tag{2.38}
\end{equation*}
$$

and it can be proved that it belongs to $\mathcal{L}\left(Y^{*}, X^{*}\right)$ with $\left\|A^{*}\right\|=\|A\|$. If $A$ is an unbounded operator, then the situation is more complicated. In general, $A^{*}$ may not exist as a single-valued operator. In other words, there may be many operators $B$ satisfying

$$
\begin{equation*}
<y^{*}, A x>=<B y^{*}, x>, \quad x \in D(A), y^{*} \in D(B) \tag{2.39}
\end{equation*}
$$

Operators $A$ and $B$ satisfying (2.39) are called adjoint to each other.
However, if $D(A)$ is dense in $X$, then there is a unique maximal operator $A^{*}$ adjoint to $A$; that is, any other $B$ such that $A$ and $B$ are adjoint to each other, must satisfy $B \subset A^{*}$. This $A^{*}$ is called the adjoint operator to $A$. It can be constructed in the following way. The domain $D\left(A^{*}\right)$ consists of all elements $y^{*}$ of $Y^{*}$ for which there exists $f^{*} \in X^{*}$ with the property

$$
\begin{equation*}
<y^{*}, A x>=<f^{*}, x> \tag{2.40}
\end{equation*}
$$

for any $x \in D(A)$. Because $D(A)$ is dense, such element $f^{*}$ can be proved to be unique and therefore we can define $A^{*} y^{*}=f^{*}$. Moreover, the assumption $\overline{D(A)}=X$ ensures that $A^{*}$ is a closed operator though not necessarily densely defined. In reflexive spaces the situation is better: if both $X$ and $Y$ are reflexive, then $A^{*}$ is closed and densely defined with

$$
\begin{equation*}
\bar{A}=\left(A^{*}\right)^{*} ; \tag{2.41}
\end{equation*}
$$

see [105, Theorems III.5.28, III.5.29].
Corollary 2.17. Let $A$ be a densely defined closed linear operator. If both $A$ and $A^{*}$ are dissipative, then $A$ is the generator of a $C_{0}$-semigroup of contractions on $X$.

Proof. It suffices to prove that, e.g., $\operatorname{Im}(I-A)=X$. Since $A$ is dissipative and closed, $\operatorname{Im}(\lambda I-A)$ is a closed subspace of $X$. Indeed, if $y_{n} \rightarrow y, y_{n} \in$ $\operatorname{Im}(I-A)$, then, by dissipativity, $\left\|x_{n}-x_{m}\right\| \leq\left\|\left(x_{n}-x_{m}\right)-\left(A x_{n}-A x_{m}\right)\right\|=$ $\left\|y_{n}-y_{m}\right\|$ and $\left(x_{n}\right)_{n \in \mathbb{N}}$ converges. But then $\left(A x_{n}\right)_{n \in \mathbb{N}}$ converges and, by closedness, $x \in D(A)$ and $x-A x=y \in \operatorname{Im}(I-A)$. Assume $\operatorname{Im}(I-A) \neq X$, then by H-B theorem, there is $0 \neq x^{*} \in X^{*}$ such that $<x^{*}, x-A x>=0$ for all $x \in D(A)$. But then $x^{*} \in D\left(A^{*}\right)$ and, by density of $D(A), x^{*}-A^{*} x^{*}=0$ but dissipativeness of $A^{*}$ gives $x^{*}=0$.

## The Cauchy problem for the heat equation

Let $C=\Omega \times(0, \infty), \Sigma=\partial \Omega \times(0, \infty)$ where $\Omega$ is an open set in $\mathbb{R}^{n}$. We consider the problem

$$
\begin{align*}
\partial_{t} u & =\Delta u, \quad \operatorname{in} \Omega \times[0, T],  \tag{2.42}\\
u & =0, \quad \text { on } \Sigma,  \tag{2.43}\\
u & =u_{0}, \quad \text { on } \Omega . \tag{2.44}
\end{align*}
$$

Theorem 2.18. Assume that $u_{0} \in L_{2}(\Omega)$ where $\Omega$ is bounded and has a $C^{2}$ boundary. Then there exists a unique function $u$ satisfying (2.44)-(1.27) such that $u \in C\left([0, \infty) ; L_{2}(\Omega)\right) \cap C\left([0, \infty) ; W_{2}^{2}(\Omega) \cap \stackrel{\circ}{W}_{2}^{1}(\Omega)\right)$,

Proof. The strategy is to consider (2.44-(1.27) as the abstract Cauchy problem

$$
u^{\prime}=A u, \quad u(0)=u_{0}
$$

in $X=L_{2}(\Omega)$ where $A$ is the unbounded operator defined by

$$
A u=\Delta u
$$

for

$$
\left.u \in D(A)=\left\{u \in \stackrel{\mathrm{o}}{W_{2}^{1}}(\Omega) ; \Delta u \in L_{2}(\Omega)\right\}=W_{2}^{2}(\Omega) \cap \stackrel{\circ}{W}_{2}^{1}(\Omega)\right)
$$

First we observe that $A$ is densely defined as $C_{0}^{\infty}(\Omega) \subset \stackrel{\circ}{W}_{2}^{1}(\Omega)$ and $\Delta C_{0}^{\infty}(\Omega) \subset$ $L_{2}(\Omega)$. Next, $A$ is dissipative. For $u \in L_{2}(\Omega), \mathcal{J} u=u$ and

$$
(A u, u)=-\int_{\Omega}|\nabla u|^{2} d \mathbf{x} \leq 0
$$

Further, we consider the variational problem associated with $I-A$, that is, to find $u \in W_{2}^{1}(\Omega)$ to

$$
a(u, v)=\int_{\Omega} \nabla u \nabla v d \mathbf{x}+\int_{\Omega} u v d \mathbf{x}=\int_{\Omega} f v d \mathbf{x}, \quad v \in \stackrel{\circ}{W}_{2}^{1}(\Omega)
$$

where $f \in L_{2}(\Omega)$ is given. Clearly, $a(u, u)=\|u\|_{1, \Omega}^{2}$ and thus is coercive. Hence there is a unique solution $u \in \stackrel{\circ}{W}{ }_{2}^{1}$ which, by writing

$$
\int_{\Omega} \nabla u \nabla v d \mathbf{x}=\int_{\Omega} f v d \mathbf{x}-\int_{\Omega} u v d \mathbf{x}=\int_{\Omega}(f-u) v d \mathbf{x}
$$

can be shown to be in $W_{2}^{2}(\Omega)$. This ends the proof of generation.
If we wanted to use the Hille-Yosida theorem instead, then to find the resolvent, we would have to solve

$$
a(u, v)=\int_{\Omega} \nabla u \nabla v d \mathbf{x}+\lambda \int_{\Omega} u v d \mathbf{x}=\int_{\Omega} f v d \mathbf{x}, \quad v \in \stackrel{\circ}{W}_{2}^{1}(\Omega)
$$

for $\lambda>0$. The procedure is the same and we get in particular for the solution

$$
\left\|\nabla u_{\lambda}\right\|_{0, \Omega}^{2}+\lambda\left\|u_{\lambda}\right\|_{0, \Omega}^{2} \leq\|f\|_{0, \Omega}\left\|u_{\lambda}\right\|_{0, \Omega}
$$

Since $u_{\lambda}=R(\lambda, A) f$ we obtain

$$
\lambda\|R(\lambda, A) f\|_{0, \Omega}^{2} \leq \lambda^{-1}\|f\|_{0, \Omega}
$$

Closedness follows from continuous invertibility.

He was able to prove that for any $x \in X, G_{\lambda}(t) x$ converges uniformly on bounded intervals as $\lambda \rightarrow \infty$ to a $C_{0}$-semigroup generated by $A$.

Another widely used approximation formula, which can also be used in the generation proof, is the operator version of the well-known scalar formula

$$
e^{a t}=\lim _{n \rightarrow \infty}\left(1-\frac{t a}{n}\right)^{-n}
$$

Precisely, [141, Theorem 1.8.3], if $A$ is the generator of a $C_{0}$-semigroup $(G(t))_{t \geq 0}$, then for any $x \in X$,

$$
\begin{equation*}
G(t) x=\lim _{n \rightarrow \infty}\left(I-\frac{t}{n} A\right)^{-n} x=\lim _{n \rightarrow \infty}\left(\frac{n}{t} R\left(\frac{n}{t}, A\right)\right)^{n} x \tag{2.45}
\end{equation*}
$$

and the limit is uniform in $t$ on bounded intervals.
Example 2.19. Suppose that $A$ generates a semigroup $\left(G_{A}(t)\right)_{t \geq 0}$ and consider $B=a A+b$, where $a>0$ and $b \in \mathbb{C}$. Then

$$
R(\lambda, B)=\frac{1}{a} R\left(\frac{\lambda-b}{a}, A\right)
$$

and the terms of the sequence in (2.45) for the operator $B$ can be written as

$$
\left(\frac{n}{t} R\left(\frac{n}{t}, B\right)\right)^{n} x=\left(\left(1+\frac{b t}{k}\right) \frac{k}{a t} R\left(\frac{k}{a t}, A\right)\right)^{k}\left(\frac{k+b t}{a t} R\left(\frac{k}{a t}, A\right)\right)^{b t} x
$$

where $k=n-b t, t>0$ fixed. Because the term $((k+b t) R(k / a t, A) / a t)^{b t} x$ converges to $x$ by (2.28), we can use Corollary 1.27 to obtain

$$
\begin{equation*}
G_{B}(t) x=\lim _{n \rightarrow \infty}\left(\frac{n}{t} R\left(\frac{n}{t}, B\right)\right)^{n} x=e^{b t} G_{A}(a t) x \tag{2.46}
\end{equation*}
$$

The semigroup $\left(G_{B}(t)\right)_{t \geq 0}$ is often referred to as the rescaled semigroup.
Remark 2.20. As we noticed earlier, a given operator $(A, D(A))$ can generate at most one $C_{0}$-semigroup. Using the Hille-Yosida theorem we can prove a stronger result which is useful later.
Proposition 2.21. Assume that the closure $(\bar{A}, D(\bar{A}))$ of an operator $(A, D)$ generates a $C_{0}$-semigroup in $X$. If $(B, D(B))$ is also a generator, such that $\left.B\right|_{D}=A$, then $(B, D(B))=(\bar{A}, D(\bar{A}))$.
Proof. Because $(B, D(B))$ is a generator, it is a closed extension of $(A, D)$. However, $(\bar{A}, D(\bar{A}))$ by definition is the smallest such extension so that $(\bar{A}, D(\bar{A})) \subset(B, D(B))$. From the Hille-Yosida theorem both operators $\lambda I-\bar{A}$ and $\lambda I-B$ are invertible for sufficiently large $\lambda$ hence, by Proposition 1.7, we obtain $B=\bar{A}$.

Without the assumption that the closure of $A$ is a generator there may be infinitely many extensions of a given operator which generate a semigroup. To see this it is enough to consider the semigroups generated by the realizations of the Laplacian subject to Dirichlet, Neumann, or mixed boundary conditions - all the generators coincide if restricted to the space of $C_{0}^{\infty}$ functions.

### 2.2.4 Standard Examples

Let us consider three relatively easy examples, variants of which appear frequently throughout the book.

Example 2.22. The maximal multiplication operator $\left(M_{a}, D\left(M_{a}\right)\right)$ was introduced in Example ??. With the function $a$ we associate the exponential $e^{t a}$. Because the exponential function $x \rightarrow e^{x}$ is continuous, the composition $e^{t a}$ is measurable on $\Omega$ for any fixed $t$. If we additionally assume

$$
\begin{equation*}
\underset{\mathbf{x} \in \Omega}{\operatorname{ess} \sup } \Re a(\mathbf{x})=\sup \left\{\Re \lambda ; \lambda \in a_{\text {ess }}(\Omega)\right\}<+\infty \tag{2.47}
\end{equation*}
$$

then $e^{t a}$ is essentially bounded on $\Omega$. We define the multiplication semigroup by

$$
\begin{equation*}
G_{a}(t) f:=e^{t a} f, \quad f \in L_{p}(\Omega), t \geq 0 \tag{2.48}
\end{equation*}
$$

Because $e^{t a} \in L_{\infty}(\Omega)$, from Example ?? we know that this is a family of bounded operators in $L_{p}(\Omega)$, having properties (i) and (ii) of Definition 2.2. To prove strong continuity, we note that $e^{t a} f \rightarrow f$ almost everywhere as $t \rightarrow 0^{+}$and, because $\left\|e^{t a}\right\|_{\infty} \leq \exp \left(t \sup \left\{\Re \lambda ; \lambda \in a_{\text {ess }}(\Omega)\right\}\right)$, we obtain

$$
\lim _{t \rightarrow 0^{+}}\left\|e^{t a} f-f\right\|_{p}=0
$$

by the dominated convergence theorem. Thus $\left(G_{a}(t)\right)_{t \geq 0}$ is a strongly continuous semigroup. It is an interesting observation, [79, Proposition I.4.12], that if $(G(t))_{t \geq 0}$ is a multiplication semigroup, that is, $G(t) f=b(t) f$ for some bounded measurable function $b$, then $b=e^{t a}$ for a measurable function $a$ satisfying $a_{\text {ess }}(\Omega)<+\infty$.

We conclude this example by showing that $\left(M_{a}, D\left(M_{a}\right)\right)$ is indeed the generator of $\left(G_{a}(t)\right)_{t \geq 0}$. Denote by $A$ the generator of $\left(G_{a}(t)\right)_{t \geq 0}$ and let $f \in D(A)$. Then

$$
\lim _{t \rightarrow 0^{+}} \frac{e^{t a} f-f}{t}=A f, \quad \text { in } L_{p}(\Omega)
$$

and there is a sequence $\left(t_{n}\right)_{n \in \mathbb{N}}$ such that

$$
\lim _{t_{n} \rightarrow 0^{+}} \frac{e^{t_{n} a} f-f}{t_{n}}=A f
$$

almost everywhere in $\Omega$. However, for almost any $\mathbf{x} \in \Omega, t \rightarrow e^{t a(\mathbf{x})} f(\mathbf{x})$ has a classical derivative at $t=0$, equal to $a(\mathbf{x}) f(\mathbf{x})$ and thus $[A f](\mathbf{x})=a(\mathbf{x}) f(\mathbf{x})$ for almost any $\mathbf{x} \in \Omega$. Thus, $D(A) \subset D\left(M_{a}\right)$. However, $\lambda I-A$ and $\lambda I-M_{a}$ are invertible for sufficiently large $\lambda$ by Theorem 2.13 and Example ?? combined with assumption (2.47), respectively. Proposition 1.7 then yields $A=M_{a}$.

Example 2.23. Let $X=L_{p}(I)$, where $I$ is either $\mathbb{R}$ or $\mathbb{R}_{+}$. In both cases we can define a (left) translation semigroup by

$$
\begin{equation*}
(G(t) f)(s):=f(t+s), \quad f \in X, \text { and } s, t \in I \tag{2.49}
\end{equation*}
$$

The semigroup property is obvious. Next, for each $t \geq 0$, we have

$$
\|G(t) f\|_{p}^{p}=\int_{I}|f(t+s)|^{p} d s \leq \int_{I}|f(r)|^{p} d r=\|f\|_{p}^{p}
$$

where, in the case $I=\mathbb{R}$, we have the equality. Hence $(G(t))_{t \geq 0}$ satisfies

$$
\begin{equation*}
\|G(t)\| \leq 1 \tag{2.50}
\end{equation*}
$$

and so $(G(t))_{t \geq 0}$ is a semigroup of contractions.
To prove that $(G(t))_{t \geq 0}$ is strongly continuous, we use an approximation approach. First let $\phi \in C_{0}^{\infty}(I)$. It is uniformly continuous (having compact support) hence for any $\epsilon>0$ there is $\delta>0$ such that for any $s \in I$ and $0<t<\delta$,

$$
|\phi(t+s)-\phi(s)|<\epsilon
$$

Thus,

$$
\int_{I}|\phi(t+s)-\phi(s)|^{p} d s \leq M_{\phi} \epsilon^{p}
$$

where $M_{\phi}$ is the measure of some fixed neighbourhood of the support of $\phi$ containing supports of all $s \rightarrow \phi(t+s)$ with $0<t<\delta$. Because $C_{0}^{\infty}(I)$ is dense in $L_{p}(I)$ for $1 \leq p<\infty$ (see Example 1.3), (2.50) allows us to use Corollary 1.28 to claim that $(G(t))_{t \geq 0}$ is a strongly continuous semigroup.

We can now use Theorem ?? to claim that there is a representation $(t, s) \rightarrow[G(t) f](s)$ of $G(t) f$ which is measurable on $\mathbb{R}_{+} \times I$ and such that the Riemann integral of $t \rightarrow G(t) f$ coincides for almost every $s \in I$ with the Lebesgue integral of $[G(t) f](s)$ with respect to $t$. Note that in this case it follows directly as the composition of a measurable function with $(t, s) \rightarrow t+s$ is measurable, [149, p. 273], but in general it is not that obvious. Hence, from now on we do not distinguish between a vector-valued function and its measurable representation.

Let us denote by $(A, D(A))$ the generator of $(G(t))_{t \geq 0}$ and let $g:=A f \in$ $L_{p}(I)$. Thus, $\Delta_{h} f:=h^{-1}(G(h) f-f) \rightarrow g$ in $L_{p}(I)$. Taking a compact interval $[a, b] \subset I$, we have

$$
\begin{aligned}
\left|\int_{a}^{b}\left(\Delta_{h} f(s)-g(s)\right) d s\right| & \leq \int_{a}^{b}\left|\Delta_{h} f(s)-g(s)\right| d s \leq|b-a|^{1 / q}\left\|\Delta_{h} f-g\right\|_{L_{p}([a, b])} \\
& \leq|b-a|^{1 / q}\left\|\Delta_{h} f-g\right\|_{L_{p}(I)}
\end{aligned}
$$

so

$$
\lim _{h \rightarrow 0^{+}} \int_{a}^{b} h^{-1}(f(s+h)-f(s)) d s=\int_{a}^{b} g(s) d s
$$

On the other hand, we can write

$$
\int_{a}^{b} h^{-1}(f(s+h)-f(s)) d s=h^{-1} \int_{b}^{b+h} f(s) d s-h^{-1} \int_{a}^{a+h} f(s) d s
$$

where the terms are the difference quotients of the function $\int_{t_{0}}^{t} f(s) d s$ at $t=a$ and $t=b$, respectively. Because $f$ is integrable on compact intervals, $\int_{t_{0}}^{t} f(s) d s \in A C(I)$ and its derivative is almost everywhere given by the integrand $f$; see Example ??. By redefining $f$ on a set of measure zero, we can write

$$
f(x)=f(a)+\int_{a}^{x} g(s) d s, \quad x \in I
$$

Thus, we see that $A \subset T$, where $T$ is the maximal differential operator on $L_{p}(I)$; see Example ??. From this example we know that $T$ is invertible, so Proposition 1.7 gives $A=T$, as in Example 2.22.

We note that the identification of the generator of the translation semigroup in Example 2.23 can be done by finding the resolvent through the Laplace transform (2.26):
$[R(\lambda, A) f](s)=\int_{0}^{\infty} e^{-\lambda t}[G(t) f](s) d t=\int_{0}^{\infty} e^{-\lambda t} f(t+s) d t=e^{\lambda s} \int_{s}^{\infty} e^{-\lambda y} f(y) d y$,
for $\lambda>0$, where the conversion of the Riemann integral of the semigroup into the Lebesgue integral follows from the discussion above. Comparing (??) with the formula above shows that $R(\lambda, A) f=R(\lambda, T) f$ for all $f \in L_{p}(I)$ and hence $A=T$.

We also note that Theorem 2.13 ensures that $T$ generates a semigroup of contractions as $T$ is closed and densely defined and estimate (??) is the same as (2.34) with $M=1$ and $\omega=0$. However, it does not provide any representation formula for the semigroup, though in this simple case one can directly solve the Cauchy problem $u_{t}^{\prime}=u_{s}^{\prime}, u(0, s)=f(s)$.

Example 2.24. The resolvent of the differential operator $T_{1}$ in $L_{p}([0,1])$ defined on the domain $D\left(T_{1}\right):=\{f \in D(T) ; f(1)=0\}$ (see Example ??) satisfies estimate (??) which gives (2.34) if $\Re \lambda>0$. Therefore $T_{1}$ is also the generator of a semigroup of contractions, say $\left(G_{T_{1}}(t)\right)_{t \geq 0}$. Considerations similar to the previous example show that it is given by

$$
\left[G_{T_{1}}(t) f\right](s)= \begin{cases}f(t+s) & \text { for } 0 \leq t+s \leq 1  \tag{2.51}\\ 0 & \text { for } t+s \geq 1\end{cases}
$$

This shows that one should be careful when looking at a semigroup generated by $A$ as the exponential $e^{t A}$ because, in this particular case, $e^{t T_{1}}$ vanishes for $t>1$.

### 2.2.5 Subspace Semigroups

There are several ways of constructing new semigroups using a given semigroup $(G(t))_{t \geq 0}$ as the starting point (see, e.g., [79, pp. 59-64]). In Example 2.19 we have already seen the so-called rescaled semigroup. In this subsection we consider a particularly important, for further applications, case of restrictions of $(G(t))_{t \geq 0}$, acting in a Banach space $X$, to a subspace $Y$ which is continuously embedded in $X$ and which is invariant under $(G(t))_{t \geq 0}$. The restriction $\left(G_{Y}(t)\right)_{t \geq 0}$ of $(G(t))_{t \geq 0}$ to $Y$ is obviously a semigroup but not necessarily a $C_{0}$-semigroup. If, however, it is strongly continuous, then we can identify the generator of $\left(G_{Y}(t)\right)_{t \geq 0}$ as the part in $Y$ of the generator $A$ of $(G(t))_{t \geq 0}$, see (1.12).
Proposition 2.25. Let $(A, D(A))$ generate a $C_{0}$-semigroup $(G(t))_{t \geq 0}$ in a Banach space $X$ and let $Y$ be a subspace continuously embedded in $X$, invariant under $(G(t))_{t \geq 0}$. If the restricted semigroup $\left(G_{Y}(t)\right)_{t \geq 0}$ is strongly continuous in $Y$ then its generator is the part $A_{Y}$ of $A$ in $Y$.

Moreover, if $Y$ is closed in $X$, then $\left(G_{Y}(t)\right)_{t \geq 0}$ is automatically strongly continuous and $A_{Y}$ is the restriction of $A$ to the domain $D(A) \cap Y$.

Proof. Denote by $(C, D(C))$ the generator of $\left(G_{Y}(t)\right)_{t \geq 0}$. Because $Y$ is continuously embedded in $X, C \subset A_{Y}$ by (2.11). To prove the reverse inclusion, let $\lambda \in \mathbb{R}$ be large enough for $R(\lambda, A)$ and $R(\lambda, C)$ to admit the integral representation (2.26):

$$
R(\lambda, C) y=\int_{0}^{\infty} e^{-\lambda t} G(t) y d t=R(\lambda, A) y, \quad y \in Y
$$

Taking $x \in D\left(A_{Y}\right)$, we obtain

$$
x=R(\lambda, A)(\lambda I-A) x=R(\lambda, C)(\lambda I-A) x \in D(C)
$$

and hence $D\left(A_{Y}\right)=D(C)$.
If $Y$ is closed, then the convergence in $Y$ is induced by the norm of $X$ and therefore $\left(G_{Y}(t)\right)_{t \geq 0}$ is strongly continuous whenever $(G(t))_{t \geq 0}$ is. Also, the limit $A y=\lim _{h \rightarrow 0^{+}} h^{-1}(G(h) y-y)$ of (2.11) exists for $y \in \bar{Y}$ if and only if $y \in D(A) \cap Y$ and hence it belongs to $Y$ by the closedness of $Y$.

In some cases the assumption that $Y$ is invariant with respect to the semigroup $(G(t))_{t \geq 0}$ can be relaxed.

Proposition 2.26. Let $B$ be a closed operator and $Y=D(B)$ be normed with the graph norm. If $A \in \mathcal{G}(M, \omega)$ generates a semigroup $(G(t))_{t \geq 0}$ and if $B$ commutes with the resolvent $R(\lambda, A)$ for some $\lambda$ with $\Re \lambda>\omega$, then $B$ commutes with $(G(t))_{t \geq 0}$ and $\left(G_{Y}(t)\right)_{t \geq 0}$ is a $C_{0}$-semigroup in $Y$ satisfying $\left\|G_{Y}(t)\right\|_{Y} \leq\|G(t)\|$.

Proof. By definition, $B$ commutes with $R(\lambda, A)$ if and only if for each $f \in$ $D(B)$ we have $R(\lambda, A) f \in D(B)$ and $B R(\lambda, A) f=R(\lambda, A) B f$; see (1.13). Thus, for any $n$ we have $B R^{n}(\lambda, A) f=R^{n}(\lambda, A) B f$. In fact, by induction we easily have that $R^{n}(\lambda, A) f \in D(B)$ for any $n \in \mathbb{N}$ provided $f \in D(B)$ and the commutativity follows by iteration. Hence for any $N \in \mathbb{N}$ and $f \in D(B)$,

$$
B \sum_{n=0}^{N}(\lambda-\mu)^{n} R(\lambda, A)^{n+1} f=\sum_{n=0}^{N}(\lambda-\mu)^{n} R(\lambda, A)^{n+1} B f
$$

Taking limits of both sides as $N \rightarrow \infty$ and using closedness of $B$ we obtain, by (2.23), that $R(\mu, A) f \in D(B)$ and $B R(\mu, A) f=R(\mu, A) B f$, provided $|\mu-\lambda|<\|R(\lambda, A)\|^{-1}$. By analytic continuation we can extend this equality to the connected component of the resolvent set $\rho(A)$ and, in particular, by (2.54) to the half-plane $\Re \lambda>\omega$. Thus, for any $t \geq 0$ and $f \in D(B)$, we obtain

$$
B\left(\frac{n}{t} R\left(\frac{n}{t}, A\right)\right)^{n} f=\left(\frac{n}{t} R\left(\frac{n}{t}, A\right)\right)^{n} B f
$$

Using again closedness of $B$ and (2.45) we see that if $f \in D(B)$, then also $\lim _{n \rightarrow \infty}(n R(n / t, A) / t)^{n} f \in D(B)$ and

$$
B G(t) f=G(t) B f
$$

It is obvious that $\left(G_{Y}(t)\right)_{t \geq 0}$ is a $C_{0}$-semigroup in $Y$ and because
$\|G(t) f\|_{D(B)}=\|G(t) f\|+\|B G(t) f\| \leq\|G(t)\|(\|f\|+\|B f\|)=\|G(t)\|\|f\|_{D(B)}$
we obtain $\left\|G_{Y}(t)\right\|_{Y} \leq\|G(t)\|$.

### 2.2.6 Sobolev Towers

We briefly describe here a somewhat related construction (see [79, pp. 124129]) which allows us to restrict semigroups to the domains $D\left(A^{n}\right)$ and, more important, extend them and their generators to larger spaces. The latter will be needed in applications to identify other extensions of generators; see Corollary ??, Lemma ??, and Corollary ??. To simplify the notation, we assume that the semigroup $(G(t))_{t \geq 0}$ generated by $A$ is of negative type so that $A^{-1} \in \mathcal{L}(X)$. This can always be achieved by rescaling the semigroup. Then, for each $n \in \mathbb{N}$, we define a new norm on $D\left(A^{n}\right)$ by

$$
\begin{equation*}
\|x\|_{n}=\left\|A^{n} x\right\| \tag{2.52}
\end{equation*}
$$

The space $X_{n}=\left(D\left(A^{n}\right),\|\cdot\|_{n}\right)$ is called the associated Sobolev space of order $n$. The introduced norm is equivalent to the graph norm due to the invertibility of $A$ so $X_{n}$ are Banach spaces. Denoting by $G_{n}(t)$ the restriction of $G(t)$ to $X_{n}$, we can prove that $\left(G_{n}(t)\right)_{t \geq 0}$ are $C_{0}$-semigroups in $X_{n}$, generated by the parts $A_{n}$ of $A$ in $X_{n}$, which are the restrictions of $A$ to $D\left(A_{n+1}\right)$.

Thus, $\left(A_{n}, D\left(A_{n}\right)\right)=\left(A, D\left(A^{n+1}\right)\right)$. We observe that each $X_{n+1}$ is densely embedded in $X_{n}$ but also, via $A_{n}$, isometrically isometric to $X_{n+1}$.

In this construction we obtained $X_{n+1}$ from $X_{n}$ but we also can invert the procedure and obtain $X_{n}$ as the completion of $X_{n+1}$ with respect to the norm

$$
\|x\|_{n}=\left\|A_{n+1}^{-1}\right\|
$$

Hence, we can construct new spaces of 'negative' order using the following recursion. Starting from $X_{0}=X$ for each $n \in \mathbb{N}$ and $X_{-n+1}$ we define

$$
\begin{equation*}
\|x\|_{-n}=\left\|A_{-n+1}^{-1} x\right\| \tag{2.53}
\end{equation*}
$$

and call the completion of $X_{-n+1}$, with respect to this norm, the associated Sobolev space of order $-n$, denoting it $X_{-n}$. The continuous (by density) extensions of the operators $G_{-n+1}(t)$ from $X_{-n+1}$ to $X_{-n}$ we denote by $G_{-n}(t)$. For example, the space $X_{-1}$ is obtained as a completion of $X$ with respect to the norm $\|x\|_{-1}=\left\|A^{-1} x\right\|$. This construction leads to the spaces and operators having properties analogous to those described above. Namely, for any $m \geq n \in \mathbb{Z}$, the following statements are valid.
(i) Each $X_{n}$ is a Banach space containing $X_{m}$ as a dense subspace.
(ii) The operators $G_{n}(t)$ form a $C_{0}$-semigroup $\left(G_{n}(t)\right)_{t \geq 0}$ on $X_{n}$.
(iii) The generator $A_{n}$ of $\left(G_{n}(t)\right)_{t \geq 0}$ has domain $D\left(A_{n}\right)=X_{n+1}$ and is the unique extension by density of $A_{m}: X_{m+1} \rightarrow X_{m}$ to an isometry from $X_{n+1}$ onto $X_{n}$.

In particular, the generator $\left(A_{-1}, X\right)$ of $\left(T_{-1}(t)\right)_{t \geq 0}$ is the unique extension by density of $(A, D(A))$.

Example 2.27. As a simple example that is useful in the sequel we consider the semigroup $(G(t))_{t \geq 0}$ on $X=X_{0}=L_{1}(\Omega, d \mu)$ generated by the multiplication operator by a function $-a$, where $a$ is assumed to be measurable and nonnegative almost everywhere on $\Omega$; see Example 2.22. Because in general $0 \in \sigma\left(M_{-a}\right)$, we use $A u=\left(I-M_{-a}\right)^{-1} u=(1+a)^{-1} u$. We have then $1+a>0$ almost everywhere on $\Omega$ and

$$
X_{n}=\left\{u \in L_{0}(\Omega, d \mu) ;(1+a)^{n} u \in L_{1}(\Omega, d \mu)\right\}, \quad n \in \mathbb{Z}
$$

Thus, in particular, $X_{-1}$ consists of these measurable functions which are integrable after multiplication by $(1+a)^{-1}$.

### 2.2.7 The Laplace Transform and the Growth Bounds of a Semigroup

It is important to note that the Hille-Yosida theorem is valid in both real and complex Banach spaces with the same formulation. Thus if $A$ is an operator
in a real Banach space $X$, generating a semigroup $(G(t))_{t \geq 0}$, then its complexification will generate a complex semigroup of the same type in the complexification $X_{C}$ of $X$ equipped with the norm (??). This allows us to extend (2.26) to complex values of $\lambda$. Precisely, [141, Remark 1.5.4], if $\Re \lambda>\omega_{0}(G)$, then $\lambda \in \rho(A)$ and

$$
\begin{equation*}
R(\lambda, A) x=\int_{0}^{\infty} e^{-\lambda t} G(t) x d t \tag{2.54}
\end{equation*}
$$

is valid for all $x \in X$. The integral in (2.54) is absolutely convergent. Moreover, iterations of the resolvent give the following formula,

$$
\begin{equation*}
R(\lambda, A)^{n} x=\frac{(-1)^{n-1}}{(n-1)!} \frac{d^{n-1}}{d \lambda^{n-1}} R(\lambda, A)=\frac{1}{(n-1)!} \int_{0}^{\infty} t^{n-1} e^{-\lambda t} G(t) x d t \tag{2.55}
\end{equation*}
$$

valid for all $x \in X$ and this yields the estimate

$$
\begin{equation*}
\left\|R(\lambda, A)^{n}\right\| \leq \frac{M}{\left(\Re \lambda-\omega_{0}(G)\right)^{n}}, \quad \Re \lambda>\omega_{0}(G) \tag{2.56}
\end{equation*}
$$

An immediate consequence of the above considerations is that the spectrum of a semigroup generator is always contained in a left half-plane. Let us recall that the location of this half-plane is given by the spectral bound

$$
\begin{equation*}
s(A)=\sup \{\Re \lambda ; \lambda \in \sigma(A)\} \tag{2.57}
\end{equation*}
$$

defined in (??). For semigroups generated by bounded operators and, in particular, by matrices, Liapunov's theorem, [112] and [79, Theorem I.2.10], states that the type $\omega_{0}(G)$ of the semigroup is equal to $s(A)$. This is no longer true for strongly continuous semigroups in general; see for example, [141, Example 4.4.2] or [136, Example A-III.1.3], where it is shown that the translation semigroup $[G(t) f](s)=f(t+s)$ on the space $X=L_{p}\left(\mathbb{R}_{+}\right) \cap E$, where $E$ is the weighted space $E:=\left\{f \in L_{p}\left(\mathbb{R}_{+}\right), e^{s} d s\right\}$, whose generator $A$ is the differentiation operator, satisfies $\omega_{0}(G)=0$ and $s(A)=-1$.

Thus at this moment we only have the obvious estimate

$$
\begin{equation*}
s(A) \leq \omega_{0}(G)<+\infty \tag{2.58}
\end{equation*}
$$

The relation between the spectral properties of the generator and the longtime behaviour of the semigroup has been a major subject of research in semigroup theory over the last several years and the results are summarized in several monographs, such as $[139,79,12]$ to mention but a few. However, most of that research does not directly pertain to the topic of the book so we shall mention just a few results of direct relevance.

That the type $\omega_{0}(G)$ might be a rather crude estimate of $s(A)$ can be expected because the former is determined by the absolute convergence of the Laplace integral and the integral converges as an improper integral in
a possibly larger half-plane $\Re \lambda>a b s(G)$; see (??). At this moment we do not know, however, whether the Laplace integral still determines there the resolvent of $A$. This question is addressed in the next proposition.

Proposition 2.28. If, for some $\lambda \in \mathbb{C}$,

$$
\begin{equation*}
B_{\lambda} x:=\lim _{\tau \rightarrow \infty} \int_{0}^{\tau} e^{-\lambda t} G(t) x d t \tag{2.59}
\end{equation*}
$$

exists for all $x \in X$, then $\lambda \in \rho(A)$ and $B_{\lambda} x=R(\lambda, A) x$ for all $x \in X$.
Proof. By replacing $G(t) x$ by $e^{-\lambda t} G(t) x$ and using Example 2.19 we can assume $\lambda=0$. Accordingly, denote $B_{0}$ by $B$. Thus

$$
\frac{1}{h}(G(h) B x-B x)=-\frac{1}{h} \int_{0}^{h} G(s) x d s \rightarrow-x
$$

by (2.13). Hence $B x \in D(A)$ and $A B x=-x$ for all $x \in X$.
Next suppose $x \in D(A)$. Then

$$
B A x=\lim _{\tau \rightarrow \infty} \int_{0}^{\tau} G(t) A x d t=\lim _{\tau \rightarrow \infty} G(\tau) x-x
$$

by (2.14) and hence $y:=\lim _{\tau \rightarrow \infty} G(\tau) x$ exists. Because $\lim _{\tau \rightarrow \infty} \int_{0}^{\tau} G(t) x d t$ also exists, we must have $y=0$ and $B A x=-x$ for any $x \in D(A)$. Because $A$ is closed, Proposition 1.33 implies $0 \in \rho(A)$ and $B=-A^{-1}=R(0, A)$.

Thus we see that $\{\lambda \in \mathbb{C} ; \Re \lambda>\operatorname{abs}(G)\} \subset \rho(A)$. It is still not clear whether $s(A)=a b s(G)$. We can prove, however, that $a b s(G)$ controls the growth of classical solutions of $(2.17),(2.18)$, that is, of the solutions emanating from $x \in D(A)$. To make this concept precise, we define the growth bound $\omega_{1}(G)$ by
$\omega_{1}(G)=\inf \left\{\omega\right.$; there is $M$ such that $\left.\|G(t) x\| \leq M e^{\omega t}\|x\|_{D(A)}, x \in D(A), t \geq 0\right\}$.
Clearly, $\omega_{1}(G) \leq \omega_{0}(G)$.
Proposition 2.29. For a semigroup $(G(t))_{t \geq 0}$ we have

$$
\begin{equation*}
\omega_{1}(G)=a b s(G) \tag{2.61}
\end{equation*}
$$

Proof. Let us fix $\omega>\omega_{1}(G)$ and take any $\lambda \in \mathbb{C}$ with $\Re \lambda>\omega$. We begin by showing that for such $\lambda$ the operator $B_{\lambda}$, defined by (2.59), exists. Let us
choose $M$ in such a way that $\|G(t) x\| \leq M e^{\omega t}\|x\|_{D(A)}$, as in (2.60). First let $x \in D(A)$. Then for any $0 \leq a \leq b$ we have

$$
\begin{aligned}
& \left\|\int_{a}^{b} e^{-\lambda t} G(t) x d t\right\| \leq \int_{a}^{b} e^{-\Re \lambda t}\|G(t) x\| d t \\
& \leq M \int_{a}^{b} e^{(\omega-\Re \lambda) t}\|x\|_{D(A)} d t=\frac{M}{\Re \lambda-\omega}\left(e^{(\omega-\Re \lambda) a}-e^{(\omega-\Re \lambda) b}\right)\|x\|_{D(A)}
\end{aligned}
$$

If $a, b \rightarrow \infty$, then the right hand converges to 0 and thus $B_{\lambda} x$ exists. Second, we consider the case $x=(\lambda I-A) y$ for some $y \in D(A)$. Because $A$ is closed and $A-\lambda I$ generates $\left(e^{-\lambda t} G(t)\right)_{t \geq 0}$ we obtain, by (2.14),

$$
\int_{0}^{\tau} e^{-\lambda t} G(t) x=(\lambda I-A) \int_{0}^{\tau} e^{-\lambda t} G(t) y d t=y-e^{-\lambda \tau} G(\tau) y
$$

and, using $y \in D(A)$ and $\Re \lambda>\omega$, we obtain

$$
\begin{equation*}
\lim _{\tau \rightarrow \infty} \int_{0}^{\tau} e^{-\lambda t} G(t) x=y-\lim _{\tau \rightarrow \infty} e^{-\lambda \tau} G(\tau) y=y \tag{2.62}
\end{equation*}
$$

Finally, from the resolvent identity (2.22), we obtain that for $x \in X$,

$$
x=(\lambda I-A) R(\mu, A) x+(\mu-\lambda) R(\mu, A) x
$$

so that any $x \in X$ can be written as a sum of elements from $D(A)$ and $\operatorname{Im}(\lambda I-A)$; thus, by the two cases considered already, $B_{\lambda} x$ exists for any $x \in X$. This shows that $\omega>a b s(G)$ and thus $a b s(G) \leq \omega_{1}(G)$.

To complete the proof we have to show $\omega_{1}(G) \leq a b s(G)$. Let $\omega>a b s(G)$ and $\Re \lambda>\omega$. Then $R(\lambda, A) x=B_{\lambda} x$ for any $x \in X$ and, because this time we know that the left-hand side of (2.62) converges to $y=R(\lambda, A) x$, we obtain

$$
\lim _{\tau \rightarrow \infty} e^{-\lambda \tau} G(\tau) R(\lambda, A) x=0
$$

and this shows $\omega \geq \omega_{1}(G)$. Therefore $a b s(G) \geq \omega_{1}(G)$.
Unfortunately, in [169] (see also [139, Example 1.2.4]), the author constructed a semigroup with $a b s(G)=\omega_{1}(G)=1$ and $s(A)=0$. Hence, in general, $s(A)$ does not provide full information about the long-time behaviour of even classical solutions. However, as we show later, for positive semigroups we have $\omega_{1}(G)=s(A)$ and for positive semigroups in $L_{p}$-spaces it is possible to prove that $s(A)=\omega_{0}(G)$.

### 2.3 Dissipative Operators

Let $X$ be a Banach space (real or complex) and $X^{*}$ be its dual. From the Hahn-Banach theorem, Theorem 1.12, and Remark ??, for every $x \in X$ there exists $x^{*} \in X^{*}$ satisfying

$$
<x^{*}, x>=\|x\|^{2}=\left\|x^{*}\right\|^{2}
$$

Therefore the duality set

$$
\begin{equation*}
\left.\mathcal{J}(x)=\left\{x^{*} \in X^{*} ;<x^{*}, x\right\rangle=\|x\|^{2}=\left\|x^{*}\right\|^{2}\right\} \tag{2.63}
\end{equation*}
$$

is nonempty for every $x \in X$.
Definition 2.30. We say that an operator $(A, D(A))$ is dissipative if for every $x \in D(A)$ there is $x^{*} \in \mathcal{J}(x)$ such that

$$
\begin{equation*}
\Re<x^{*}, A x>\leq 0 . \tag{2.64}
\end{equation*}
$$

If $X$ is a real space, then the real part in the above definition can be dropped. An important equivalent characterisation of dissipative operators, [141, Theorem 1.4.2], is that $A$ is dissipative if and only if for all $\lambda>0$ and $x \in D(A)$,

$$
\begin{equation*}
\|(\lambda I-A) x\| \geq \lambda\|x\| \tag{2.65}
\end{equation*}
$$

We note some important properties of dissipative operators.
Proposition 2.31. [79, Proposition II.3.14] If $(A, D(A))$ is dissipative, then
(i) $\lambda I-A$ is one-to-one for any $\lambda>0$ and

$$
\begin{equation*}
\left\|(\lambda I-A)^{-1} x\right\| \leq \frac{1}{\lambda}\|x\| \tag{2.66}
\end{equation*}
$$

for all $x \in \operatorname{Im}(\lambda I-A)$.
(ii) $\operatorname{Im}(\lambda I-A)=X$ for some $\lambda>0$ if and only if $\operatorname{Im}(\lambda I-A)=X$ for all $\lambda>0$.
(iii) $A$ is closed if and only if $\operatorname{Im}(\lambda I-A)$ is closed for some (and hence all) $\lambda>0$.
(iv) If $A$ is densely defined, then $A$ is closable and $\bar{A}$ is dissipative. Moreover, $\overline{\operatorname{Im}(\lambda I-A)}=\operatorname{Im}(\lambda I-\bar{A})$.

Combination of the Hille-Yosida theorem with the above properties gives a generation theorem for dissipative operators, known as the Lumer-Phillips theorem ([141, Theorem 1.43] or [79, Theorem II.3.15]).

Theorem 2.32. For a densely defined dissipative operator $(A, D(A))$ on a Banach space $X$, the following statements are equivalent.
(a) The closure $\bar{A}$ generates a semigroup of contractions.
(b) $\overline{\operatorname{Im}(\lambda I-A)}=X$ for some (and hence all) $\lambda>0$.

If either condition is satisfied, then A satisfies (2.64) for any $x^{*} \in \mathcal{J}(x)$.
In particular, if we know that $A$ is closed then the density of $\operatorname{Im}(\lambda I-A)$ is sufficient for $A$ to be a generator. On the other hand, if we do not know a priori that $A$ is closed then $\operatorname{Im}(\lambda I-A)=X$ yields $A$ being closed and consequently that it is the generator.

Example 2.33. The multiplication semigroup of Example 2.48 is a semigroup of contractions only if $a_{e s s}(\Omega) \leq 0$.

The maximal differential operator $T$ on $L_{p}(I), 1 \leq p<\infty$, where $I=\mathbb{R}$ or $I=\mathbb{R}_{+}$, discussed in Example ??, is densely defined $\left(C_{0}^{\infty}(I) \subset D(T)\right)$ and dissipative by (??) and (2.65). Thus the translation semigroups are semigroups of contractions, which was proved directly in Example 2.23.

Also the differential operator $T_{1}$ of Example 2.24 is densely defined and dissipative by (??). Hence it generates a semigroup of contractions in $L_{p}([0,1])$, $1 \leq p<\infty$. An interesting feature of this operator is discussed in Example 2.35 below.

We now provide a few variations of the Lumer-Phillips theorem.
Example 2.34. If $(A, D(A))$ is a densely defined operator in $X$ and both $A$ and its adjoint $A^{*}$ are dissipative, then $\bar{A}$ generates a semigroup of contractions in $X$. In fact, because $\bar{A}$ is dissipative and closed, $\operatorname{Im}(I-\bar{A})$ is closed. If $\operatorname{Im}(I-\bar{A}) \neq X$, then for some $0 \neq x^{*} \in X^{*}$ we have

$$
0=<x^{*}, x-\bar{A} x>=<x^{*}-\bar{A}^{*} x^{*}, x>
$$

for all $x \in D(\bar{A})$. Because $\bar{A}$ is densely defined, $x^{*}-\bar{A}^{*} x^{*}=0$ and because $\bar{A}^{*}$ is dissipative, $x^{*}=0$. Hence $\operatorname{Im}(I-\bar{A})=X$ and $\bar{A}$ is the generator of a dissipative semigroup by Theorem 2.32. In particular, dissipative self-adjoint operators on Hilbert spaces are always generators.

Example 2.35. The assumption of the density of $D(A)$ can be circumvented to a certain extent. If $(A, D(A))$ is a dissipative operator in $X$ with $\operatorname{Im}(\lambda I-A)=$ $X$ for some $\lambda>0$ and possibly $\overline{D(A)} \neq X$, then the part of $A$ in $X_{0}=$ $\overline{D(A)}$ (see (1.12)) is densely defined in $X_{0}$ and generates there a semigroup of contractions.

A classical example in such a case is the realisation of the differential operator $T_{1}$ in the space of continuous functions. In fact, define $A f=f^{\prime}$ on the domain $D(A)=\left\{f \in C^{1}([0,1]) ; f(1)=0\right\}$ in $X=C([0,1]) . A$ is a closed, dissipative, and surjective operator but $D(A)$ is not dense. However, restricted to the domain $\left\{f \in C^{1}([0,1]) ; f(1)=0, f^{\prime}(1)=0\right\}$, $A$ generates a semigroup of contractions in $X_{0}=\{f \in C([0,1]) ; f(1)=0\}$. The semigroup is again given by the left translation (2.51). However, it cannot be extended to the whole $X$ as it would not give a continuous function if $f(1) \neq 0$.

The situation described in the previous example cannot occur in reflexive spaces. Precisely speaking, [141, Theorem 1.4.6], if $(A, D(A))$ is a dissipative operator on a reflexive Banach space $X$, such that $\operatorname{Im}(\lambda I-A)=X$ for some $\lambda>0$, then it is densely defined.

### 2.3.1 Application: Diffusion Problems

Some of the most important examples of contractive semigroups which occur in applications are those describing diffusion processes. Their theory has been very well developed but is rather tangential to the subject studied in this book so we discuss them rather briefly, focusing only on those aspects that are needed later. Unfortunately, even such a superficial survey requires a substantial theoretical machinery. A more comprehensive account of various aspects of the theory can be found in [71, Chapter 1], [82, Chapter 4 ], and [79, Section VI.5] among others. We begin with basic definitions and facts from the Sobolev space theory (see, e.g., [4, 93]).

Let $\Omega$ be an open subset of $\mathbb{R}^{n}$, possibly equal to the whole space. Recall that Sobolev spaces $W_{p}^{m}(\Omega), 1 \leq p \leq \infty, m \in \mathbb{N}_{0}$, are defined as

$$
\begin{equation*}
W_{p}^{m}(\Omega):=\left\{u \in L_{p}(\Omega) ; \partial^{\alpha} u \in L_{p}(\Omega), 0 \leq|\alpha| \leq m\right\} \tag{2.67}
\end{equation*}
$$

where $\partial^{\alpha}=\partial_{x_{1}}^{\alpha_{1}} \ldots \partial_{x_{n}}^{\alpha_{n}},|\alpha|=\alpha_{1}+\ldots \alpha_{n}$ is the distributional derivative of order $|\alpha|$, introduced in Example ??. Endowed with the norm

$$
\begin{equation*}
\|u\|_{m, p}:=\|u\|_{W_{p}^{m}(\Omega)}:=\left(\sum_{0 \leq|\alpha| \leq m}\left\|\partial^{\alpha} u\right\|_{L_{p}(\Omega)}^{p}\right)^{1 / p} \tag{2.68}
\end{equation*}
$$

the space $W_{p}^{m}(\Omega)$ becomes a Banach space. In the particular case of $p=2$, (2.68) defines a Hilbert space norm with the corresponding scalar product given by

$$
(u, v)_{W_{2}^{m}(\Omega)}:=\sum_{0 \leq|\alpha| \leq m} \int_{\Omega} \partial^{\alpha} u(\mathbf{x}) \overline{\partial^{\alpha} v(\mathbf{x})} d \mathbf{x}
$$

The space $C_{0}^{\infty}(\Omega)$ is continuously embedded in any $W_{p}^{m}(\Omega)$ but the embedding is not dense unless $\Omega=\mathbb{R}^{n}$ (or $m=0$ ). However, the closure of $C_{0}^{\infty}(\Omega)$ in the $W_{p}^{m}(\Omega)$-norm, denoted as

$$
\stackrel{o}{W}_{p}^{m}(\Omega)={\overline{C_{0}^{\infty}(\Omega)}}_{W_{p}^{m}(\Omega)}
$$

is very important in applications through its connection with boundary values of functions from $W_{p}^{m}(\Omega)$.

It is possible to extend the definition of Sobolev spaces to fractional orders by defining $W_{p}^{r}(\Omega)$, where $r=m+\sigma, m$ is an integer, and $0<\sigma<1$, by requiring that

$$
\int_{\Omega} \int_{\Omega} \frac{\left|\partial^{\alpha} u(\mathbf{x})-\partial^{\alpha} u(\mathbf{y})\right|^{p}}{|\mathbf{x}-\mathbf{y}|^{n+\sigma p}} d \mathbf{x} d \mathbf{y}<+\infty
$$

for all $|\alpha|=m$ (see, e.g., [93, Definition 1.3.2.1]) but we make little use of these spaces and we therefore do not enter into details of their theory.

In what follows we assume that if the boundary $\partial \Omega$ of $\Omega$ is not an empty set, then $\partial \Omega$ is an $n-1$ dimensional manifold of class $C^{k, 1}, k \geq 0$ (that is, the local atlas of $\partial \Omega$ is $k$ times continuously differentiable with the derivatives of order $k$ being Lipschitz continuous). For a smooth function $u$ on $\Omega$, let us define the trace of $u$ on $\partial \Omega$ to be the pointwise restriction of $u$ to $\partial \Omega$ :

$$
\gamma u=\left.u\right|_{\partial \Omega} .
$$

If $m>1 / p$ is not an integer, $m \leq k+1$ and $l+\sigma=m-1 / p, 0<\sigma<1$, $l \geq 0$ an integer, then the mapping

$$
u \rightarrow\left(\gamma u, \gamma \frac{\partial u}{\partial \nu}, \ldots, \gamma \frac{\partial^{l} u}{\partial \nu^{l}}\right)
$$

where $\partial / \partial \nu$ denotes the outward normal derivative at $\partial \Omega$, can be extended by density to a continuous mapping from $W_{p}^{m}(\Omega)$ to $\left(L_{2}(\partial \Omega)\right)^{l+1}$ (precisely speaking onto a product of appropriate Sobolev spaces of fractional order defined on $\partial \Omega)$.

Under these assumptions we have another characterisation of $\stackrel{\circ}{W}_{p}^{m}(\Omega)$ :

$$
\begin{equation*}
\stackrel{\circ}{W}_{p}^{m}(\Omega)=\left\{u \in W_{p}^{m}(\Omega) ; \gamma u=\gamma \frac{\partial u}{\partial \nu}=\cdots=\gamma \frac{\partial^{l} u}{\partial \nu^{l}}=0\right\} . \tag{2.69}
\end{equation*}
$$

Let us consider the Cauchy problem of a diffusion type:

$$
\begin{align*}
\partial_{t} u(t, \mathbf{x}) & =\sum_{i, j=1}^{n} \partial_{x_{i}}\left(a_{i j}(\mathbf{x}) \partial_{x_{j}} u(t, \mathbf{x})\right) \\
u(0, \mathbf{x}) & =u_{0}(\mathbf{x}) \tag{2.70}
\end{align*}
$$

where $t>0, \mathbf{x} \in \Omega$, and $u_{0}$ is a given function. The real coefficients $\left\{a_{i j}(\mathbf{x})\right\}_{1 \leq i, j \leq n}$ are supposed to satisfy $a_{i j} \in W_{\infty}^{1}(\Omega)$ and $a_{i j}=a_{j i}$ for $i, j=1, \ldots, n$. If $\partial \Omega \neq \emptyset$, then the problem (2.70) should be supplemented by some boundary conditions defined on $\partial \Omega$; here we confine ourselves to the homogeneous Dirichlet problem; that is, we require

$$
\begin{equation*}
\left.u\right|_{\partial \Omega}=0 \tag{2.71}
\end{equation*}
$$

According to our general philosophy, we convert (2.70) and (2.71) into an abstract Cauchy problem in the Banach space $X=L_{p}(\Omega), 1 \leq p<\infty$. Our main interest is $p=1$, but most results are based on the $L_{2}$ theory so we discuss the latter setting in some detail. Let us denote by $A_{0}$ the differential expression

$$
\begin{equation*}
(A u)(\mathbf{x}):=\sum_{i, j=1}^{n} \partial_{x_{i}}\left(a_{i j}(\mathbf{x}) \partial_{x_{j}} u(\mathbf{x})\right) \tag{2.72}
\end{equation*}
$$

understood, if necessary, in the sense of distributions, and define

$$
A_{0, p} u=A u
$$

for

$$
u \in D:=C_{0}^{2}(\Omega)
$$

where $C_{0}^{2}(\Omega)$ is the space of twice-differentiable compactly supported functions in $\Omega$. The index $p$ indicates that $A_{0, p}$ is considered in the space $L_{p}(\Omega)$.

A crucial assumption is that $A$ is strongly elliptic in $\Omega$; that is, for some constant $c>0$ and all $\mathbf{y}=\left(y_{1}, \ldots, y_{n}\right) \in \mathbb{R}^{n}$ and all $\mathbf{x} \in \bar{\Omega}$ we have

$$
\begin{equation*}
\sum_{i, j=1}^{n} a_{i j}(\mathbf{x}) y_{i} y_{j} \geq c|\mathbf{y}|^{2} \tag{2.73}
\end{equation*}
$$

or equivalently

$$
\Re \sum_{i, j=1}^{n} a_{i j}(\mathbf{x}) z_{i} \bar{z}_{j} \geq c|\mathbf{z}|^{2}
$$

for all $\mathbf{z} \in \mathbb{C}^{n}$. Then we have the following result, [82, Lemma 4.4.3].
Lemma 2.36. The operator $\left(A_{0, p}, D\right)$ is dissipative for any $p \in[1, \infty]$.
By Proposition 2.31 (iv), $\left(A_{0, p}, D\right)$ is closable with dissipative closure. However, finding the $m$-dissipative extension of $A_{0, p}$ requires a deep theory.

## $L_{2}$ Theory

Let $\Omega$ be either $\mathbb{R}^{n}$ or an open set with a $C^{0,1}$ boundary $\partial \Omega$. Possibly the easiest approach to proving solvability of (2.70) in the space $L_{2}(\Omega)$ is to use the variational approach and look for a suitable realisation of $A_{0, p}$ via the associated sesquilinear form

$$
\begin{equation*}
a(u, v):=\int_{\Omega} \sum_{i, j=1}^{n} a_{i j}(\mathbf{x}) \partial_{x_{i}} u(\mathbf{x}) \partial_{x_{j}} \overline{v(\mathbf{x})} d \mathbf{x} \tag{2.74}
\end{equation*}
$$

for $u, v \in D(a)=\stackrel{\circ}{W}_{2}^{1}(\Omega)$. Note that for $\Omega=\mathbb{R}^{n}$, we have ${ }_{W}^{\circ}{ }_{2}^{1}\left(\mathbb{R}^{n}\right)=W_{2}^{1}\left(\mathbb{R}^{n}\right)$ so that we can use common notation for both spaces without causing any confusion. If $\partial \Omega \neq \emptyset$, then ${ }_{W}^{\mathrm{o}}{ }_{2}^{1}$ consists of those $W_{2}^{1}(\Omega)$ functions whose trace on $\partial \Omega$ is zero.

Then (see, e.g., [79, Theorem VI.5.18] or [82, Theorem 4.6.6]) we have:
Theorem 2.37. There is a unique dissipative operator $\left(-A_{2}, D\left(A_{2}\right)\right)$ such that $D\left(A_{2}\right) \subset D(a)$ and $a(u, v)=\left(A_{2} u, v\right)_{L_{2}(\Omega)}$ for all $u \in D\left(A_{2}\right)$ and $v \in$ $L_{2}(\Omega)$. The operator $\left(-A_{2}, D\left(A_{2}\right)\right)$ generates a semigroup of contractions in $L_{2}(\Omega)$, denoted by $\left(G_{A_{2}}(t)\right)_{t \geq 0}$.

By restricting

$$
a(u, v)=\left(A_{2} u, v\right)
$$

to $v \in C_{0}^{\infty}(\Omega)$, it is easy to see that $A_{2}$ coincides with the expression $A$ in the distributional sense; thus $D\left(A_{2}\right)$ can be characterised as

$$
D\left(A_{2}\right)=\left\{u \in \stackrel{\circ}{W}_{2}^{1}(\Omega) ; A u \in L_{2}(\Omega)\right\}
$$

(see, e.g., [93, Theorem 2.2.1.2]). The fact that $u$ satisfies the boundary condition $\gamma u=0$ if $\partial \Omega \neq \emptyset$ follows from the fact that $D\left(A_{2}\right) \subset D(a)={ }^{\circ}{ }_{2}^{1}(\Omega)$.

This result is not fully satisfactory. First, the property $A_{0} u \in L_{2}(\Omega)$ does not ensure that the second derivatives of $u$ are in $L_{2}(\Omega)$ - there may be a cancellation of singularities in the expression $A$. Second, $A_{2}$ may fall short of the 'maximal' operator $A_{2, \max }$ (see Section 2.7 and [93, p.54]) defined on

$$
D\left(A_{2, \max }\right):=\left\{u \in L_{2}(\Omega) ; A u \in L_{2}(\Omega), \gamma u=0\right\}
$$

where the trace of $u \in L_{2}(\Omega)$ such that $A u \in L_{2}(\Omega)$ can be defined by means of Green's formula; see [93, p.54].

The first question is addressed by proving that, under certain assumptions, the variational solution $u \in D\left(A_{2}\right)$ is in $W_{2}^{2}(\Omega)$. We can state the following result (see, e.g., [93, Theorems 2.2.2.3, 2.5.2.1, and 3.2.1.2], [79, VI.5.22]).

Theorem 2.38. If $\Omega=\mathbb{R}^{n}$, or $\Omega$ is convex, or $\partial \Omega$ is of class $C^{1,1}$, then

$$
\begin{equation*}
D\left(A_{2}\right)=\stackrel{\mathrm{o}}{W_{2}^{1}}(\Omega) \cap W_{2}^{2}(\Omega) \tag{2.75}
\end{equation*}
$$

In all these cases we also have

$$
\begin{equation*}
D\left(A_{2, \max }\right)=\stackrel{\mathrm{o}}{W_{2}^{1}}(\Omega) \cap W_{2}^{2}(\Omega) \tag{2.76}
\end{equation*}
$$

It is known that if $\Omega$, for instance, is a nonconvex polygon in $\mathbb{R}^{2}$, then $D\left(A_{2}\right) \neq \stackrel{\circ}{W}{ }_{2}^{1}(\Omega) \cap W_{2}^{2}(\Omega),[93$, Chapter 4$]$, and consequently we have the sequence of strict inclusions (see [20])

$$
\begin{equation*}
\stackrel{\circ}{W}_{2}^{1}(\Omega) \cap W_{2}^{2}(\Omega) \varsubsetneqq D\left(A_{2}\right) \varsubsetneqq D\left(A_{2, \max }\right) \tag{2.77}
\end{equation*}
$$

We explore some other consequences of this fact in Example 2.66.
For further reference we note the following result.
Corollary 2.39. If $B$ is a generator of a semigroup in $L_{2}\left(\mathbb{R}^{n}\right)$ and satisfies $\left.B\right|_{C_{0}^{\infty}\left(\mathbb{R}^{n}\right)}=\left.A\right|_{C_{0}^{\infty}\left(\mathbb{R}^{n}\right)}$, then $B=A_{2}$.

Proof. By (2.76) and (2.75), the graph norm generated by $A$ on $L_{2}\left(\mathbb{R}^{n}\right)$ is equivalent to the $W_{2}^{2}\left(\mathbb{R}^{n}\right)$ norm. Because $C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ is dense in $W_{2}^{2}\left(\mathbb{R}^{n}\right)$, it is a core of $A_{2}$ and the thesis follows from Proposition 2.21.

If $\partial \Omega \neq \emptyset$, then $C_{0}^{\infty}(\Omega)$ is not dense in $\stackrel{\circ}{W}_{2}^{1}(\Omega) \cap W_{2}^{2}(\Omega)$ (the closure is $\left.{ }^{\mathrm{o}}{ }_{2}^{2}(\Omega)\right)$ and we cannot expect such a result in this case. In fact, as noted in Remark 2.20, an elliptic operator with various boundary conditions generates different semigroups and yet it is given by the same expression on $C_{0}^{\infty}(\Omega)$.

We also note that by [71, Proposition 1.3.5 and Theorem 1.3.2], the resolvent $R\left(\lambda, A_{2}\right)$ is a positive operator and therefore $\left(G_{A_{2}}(t)\right)_{t \geq 0}$ is a positive semigroup (see Section 2.5).

## $L_{p}$ Theory, $p \neq 2$

We can construct $L_{p}$ realizations of $\left(G_{A_{2}}(t)\right)_{t \geq 0}$ in a relatively straightforward manner using the following theorem, [71, Theorem 1.4.1].

Theorem 2.40. Let $\left(G_{A_{2}}(t)\right)_{t \geq 0}$ be the semigroup constructed in Theorem 2.37. Then $L_{1}(\Omega) \cap L_{\infty}(\Omega)$ is invariant under $\left(G_{A_{2}}(t)\right)_{t \geq 0}$ and $\left(G_{A_{2}}(t)\right)_{t \geq 0}$ can be extended from $L_{1}(\Omega) \cap L_{\infty}(\Omega)$ to a positive one-parameter semigroup $\left(G_{p}(t)\right)_{t \geq 0}$ on $L_{p}(\Omega)$ for any $p \in[1, \infty]$. These semigroups are strongly continuous for $p \in[1, \infty)$, and are consistent in the sense that

$$
\begin{equation*}
G_{p}(t) f=G_{q}(t) f, \quad t \geq 0 \tag{2.78}
\end{equation*}
$$

for any $f \in L_{p}(\Omega) \cap L_{q}(\Omega)$.
Denoting by $A_{p}$ the generator of $\left(G_{p}(t)\right)_{t \geq 0}, 1 \leq p<+\infty$, we also have

$$
\begin{equation*}
A_{p} u=A_{q} u, \quad u \in D\left(A_{p}\right) \cap D\left(A_{q}\right) \tag{2.79}
\end{equation*}
$$

This theorem, although settling the question of the existence of semigroups in $L_{p}$ spaces, is not entirely satisfactory because it does not provide a full characterisation of generators. It can be proved that for $1<p<+\infty$ the situation parallels the $L_{2}$ case. Classical results (see, e.g., [93, Theorem 2.4.2.4] or [82, Theorem 4.8.3 and Corollary 4.8.10]) state that if $\Omega$ is a bounded domain with sufficiently smooth boundary, then $A_{p}$ is the closure in $L_{p}(\Omega)$ of $\left(A, D_{0}\right)$, where $D_{0}$ is the set of $C^{2}$ functions satisfying the homogeneous Dirichlet boundary condition on $\partial \Omega$ and

$$
\begin{equation*}
D\left(A_{p}\right)=\stackrel{\circ}{W}_{p}^{1}(\Omega) \cap W_{p}^{2}(\Omega) \tag{2.80}
\end{equation*}
$$

One can prove that also in this case $A_{p}$ coincides with the maximal operator.
The $L_{1}$ case is much more delicate. For bounded domains one can prove, [82, Theorems 4.8.3 and 4.8.17] and [62, Theorem 8], that

$$
A_{1}={\overline{\left(A, D_{0}\right)}}^{L_{1}(\Omega)}
$$

and

$$
D\left(A_{1}\right)=\left\{u \in L_{1}(\Omega) ; A_{0} u \in L_{1}(\Omega)\right\}
$$

so that $A_{1}$ is the maximal operator. However, it is no longer true that $D\left(A_{1}\right) \subset$ $W_{1}^{2}(\Omega)$ so that the second derivatives are no longer integrable: they have 'nearly' this property as $D\left(A_{1}\right) \subset \stackrel{\circ}{W}_{1}^{r}(\Omega)$ for any $r<n /(n-1)$.

Our main interest lies with the problem (2.70) in $L_{1}\left(\mathbb{R}^{n}\right)$. In this case the theory is also quite involved and the characterisation of the domain of generators is still a subject of ongoing research (see, e.g., [115]). Contrary to the case of bounded domains, the Sobolev spaces $W_{1}^{k}\left(\mathbb{R}^{n}\right)$ are not really suitable here, but often we can get an alternative characterisation using other types of spaces. We describe the case when the differential expression $A$ is the Laplacian:

$$
A u=\Delta u
$$

understood in the sense of distributions. It is useful to denote $A_{0}=\left.A\right|_{C_{0}^{\infty}\left(\mathbb{R}^{n}\right)}$. It can be shown that the realisation $A_{1}$ of the Laplacian that generates a semigroup in $L_{1}\left(\mathbb{R}^{n}\right)$ is the restriction of $A$ to the Bessel potential space defined via Fourier transform $\mathcal{F}$ as

$$
\begin{equation*}
L_{1,2}\left(\mathbb{R}^{n}\right):=\left\{u \in L_{1}\left(\mathbb{R}^{n}\right) ; \mathcal{F}^{-1}[w[\mathcal{F} u]] \in L_{1}\left(\mathbb{R}^{n}\right)\right\} \tag{2.81}
\end{equation*}
$$

where $w(y):=\left(1+|y|^{2}\right)$ (see, e.g., [98], pp. 32-37). Note that $w$ is the Fourier transform of the distributional operator $I-A$ hence, in particular, if $u \in$ $L_{1,2}\left(\mathbb{R}^{n}\right)$, then $A u \in L_{1}\left(\mathbb{R}^{n}\right)$. We norm this space with

$$
\begin{equation*}
\|u\|_{1,2}=\left\|\mathcal{F}^{-1}[w \mathcal{F}[u]]\right\|_{L_{1}\left(\mathbb{R}^{n}\right)}=\|(I-A) u\|_{L_{1}\left(\mathbb{R}^{n}\right)} . \tag{2.82}
\end{equation*}
$$

One can show that $C_{0}^{\infty}\left(\mathbb{R}^{n}\right)$ is dense in $L_{1,2}\left(\mathbb{R}^{n}\right)$ (see, e.g., [4], p. 221) so that $A_{1}={\overline{A_{0}}}^{L_{1}\left(\mathbb{R}^{n}\right)}$, as when $\Omega$ is bounded. One can also prove (see, e.g., [115]) that

$$
\begin{equation*}
L_{1,2}\left(\mathbb{R}^{n}\right) \subset W_{1}^{r}\left(\mathbb{R}^{n}\right), \quad r<2 \tag{2.83}
\end{equation*}
$$

Clearly, also $W_{2}^{2}\left(\mathbb{R}^{n}\right) \subset L_{1,2}\left(\mathbb{R}^{n}\right)$.
The semigroup generated by the Laplacian in $L_{1}\left(\mathbb{R}^{n}\right)$ is given by the classical convolution formula

$$
\begin{equation*}
[G(t) f](\mathbf{x})=\left[\mu_{t} * f\right](\mathbf{x})=\int_{\mathbb{R}^{n}} \mu_{t}(\mathbf{x}-\mathbf{y}) f(\mathbf{y}) d \mathbf{y} \tag{2.84}
\end{equation*}
$$

of the fundamental solution $\mu_{t}(\mathbf{x})=(4 \pi t)^{-n / 2} e^{-|\mathbf{x}|^{2} / 4 t}$ and the initial data $f$ ([79, p. 69] or [98, p. 32-37]). Inasmuch as

$$
\begin{equation*}
\partial_{\mathbf{x}}^{\beta}\left[\mu_{t} * f\right](\mathbf{x})=\left[\mu_{t} * \partial_{\mathbf{x}}^{\beta} f\right](\mathbf{x}) \tag{2.85}
\end{equation*}
$$

in the sense of distributions, from the Young inequality (??) with $p=q=$ $r=1$, we immediately get that $(G(t))_{t \geq 0}$ is a strongly continuous semigroup in any $W_{1}^{l}\left(\mathbb{R}^{3}\right)$, by Proposition 2.26. In particular, for $l=0$ we clearly have

$$
\|G(t) f\|_{L_{1}\left(\mathbb{R}^{n}\right)} \leq\left\|\mu_{t}\right\|_{L_{1}\left(\mathbb{R}^{n}\right)}\|f\|_{L_{1}\left(\mathbb{R}^{n}\right)}=\|f\|_{L_{1}\left(\mathbb{R}^{n}\right)}
$$

In Section ?? we need the scale of Bessel potential spaces $L_{1, s}\left(\mathbb{R}^{n}\right)$ defined in a natural way as

$$
\begin{equation*}
L_{1, s}\left(\mathbb{R}^{n}\right):=\left\{u \in L_{1}\left(\mathbb{R}^{n}\right) ; F^{-1}\left[w^{s / 2} F[u]\right] \in L_{1}\left(\mathbb{R}^{n}\right)\right\} \tag{2.86}
\end{equation*}
$$

with norms given analogously to (2.82). It can be proved that these spaces coincide with the domains of fractional powers of the operator $I-A_{1}$. Therefore the moment inequality (e.g., [141, p. 73] or [98, p. 37]) is valid: for any nonnegative $\beta \geq s \geq \gamma$ and $\theta \in[0,1]$ satisfying $s=\theta \beta+(1-\theta) \gamma$ and some constant $C>0$ we have

$$
\begin{equation*}
\|u\|_{L_{1, s}\left(\mathbb{R}^{n}\right)} \leq C\|u\|_{L_{1, \beta}\left(\mathbb{R}^{n}\right)}^{\theta}\|u\|_{L_{1, \gamma}\left(\mathbb{R}^{n}\right)}^{1-\theta}, \quad u \in L_{1, \beta}\left(\mathbb{R}^{n}\right) \tag{2.87}
\end{equation*}
$$

Applying Hölder's inequality to (2.87) with particular values $\beta=2, \gamma=0$, we obtain the second moment inequality

$$
\begin{equation*}
\|u\|_{L_{1, s}\left(\mathbb{R}^{n}\right)} \leq K\left(\epsilon\|u\|_{L_{1,2}\left(\mathbb{R}^{n}\right)}+\epsilon^{s / s-2}\|u\|_{L_{1}\left(\mathbb{R}^{n}\right)}\right) \tag{2.88}
\end{equation*}
$$

valid for some constant $K>0$, any $\epsilon>0$, and any $u \in L_{1,2}\left(\mathbb{R}^{n}\right)$. Inequality (2.88) is of importance in Theorem ??.

Unfortunately, the Bessel potential spaces $L_{1, s}\left(\mathbb{R}^{n}\right)$ do not coincide with Sobolev spaces unless $n=1$, but on the other hand, they are 'close' to them ([98], p. 35-36). In particular, we have

$$
\begin{equation*}
L_{1, s^{\prime}}\left(\mathbb{R}^{n}\right) \subset W_{1}^{1}\left(\mathbb{R}^{n}\right) \subset L_{1, s^{\prime \prime}}\left(\mathbb{R}^{n}\right) \quad \text { for } s^{\prime \prime}<1<s^{\prime} \tag{2.89}
\end{equation*}
$$

where all embeddings are continuous.
This result, combined with (2.88), allows us to treat diffusion problems with convection (represented by a suitable first-order term) by means of perturbation techniques; see Section ??.

It is important to realize that the space $L_{1,2}(\mathbb{R})$ is only practically suitable for operators having the Laplacian as their principal part because even a linear change of variables changes the domain of the generator. In fact, consider, for instance, the generator $B$ being the realisation of the expression of $u_{x_{1} x_{1}}+$ $2 u_{x_{2} x_{2}}$ in $L_{1}\left(\mathbb{R}^{2}\right)$. If $u \in D(B)=L_{1,2}\left(\mathbb{R}^{2}\right)$, then we would have $u_{x_{2} x_{2}} \in$ $L_{1}\left(\mathbb{R}^{n}\right)$, as $\Delta u \in L_{1}\left(\mathbb{R}^{n}\right)$ but then also $u_{x_{1} x_{1}} \in L_{1}\left(\mathbb{R}^{n}\right)$ and consequently we would have $L_{1,2}\left(\mathbb{R}^{2}\right)=W_{1}^{2}\left(\mathbb{R}^{2}\right)$, which is false; see [115].

### 2.3.2 Contractive Semigroups with a Parameter

In many instances we are given a family of generators depending on a parameter. It is a natural question as to whether we can patch these generators in such a way that the obtained object is again a generator in a product space. If the generators are dissipative, then the result is positive, as follows from the proposition below.

Let us consider the space $\mathcal{X}:=L_{p}(\Omega, X)$, where $1 \leq p<\infty,(\Omega, \mu)$ is a measure space and $X$ is a Banach space. Let us suppose that we are given a family of operators $\left\{\left(A_{v}, D\left(A_{v}\right)\right)\right\}_{v \in \Omega}$ in $X$ and define the operator $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ acting in $\mathcal{X}$ according to the following formulae,

$$
\begin{equation*}
\mathcal{D}(\mathcal{A}):=\left\{u \in \mathcal{X} ; u(v) \in D\left(A_{v}\right) \text { for a.e. } v \in \Omega, \mathcal{A} u \in \mathcal{X}\right\} \tag{2.90}
\end{equation*}
$$

and, for $u \in \mathcal{D}(\mathcal{A})$,

$$
\begin{equation*}
(\mathcal{A} u)(v):=A_{v} u(v) \tag{2.91}
\end{equation*}
$$

for almost every $v \in \Omega$. We have the following proposition.
Proposition 2.41. If $A_{v}$ are m-dissipative operators in $X$ for any $v \in \Omega$ and the function $v \rightarrow R\left(\lambda, A_{v}\right) f(v)$ is measurable for any $\lambda>0$ and $f \in \mathcal{X}$, then the operator $\mathcal{A}$ is an m-dissipative operator in $\mathcal{X}$. If $\left(G_{v}(t)\right)_{t \geq 0}$ and $(\mathcal{G}(t))_{t \geq 0}$ are semigroups generated by $A_{v}$ and $\mathcal{A}$, respectively, then for almost every $v \in \Omega, t \geq 0$, and $u \in \mathcal{X}$ we have

$$
\begin{equation*}
[\mathcal{G}(t) u](v)=G_{v}(t) u(v) \tag{2.92}
\end{equation*}
$$

Proof. Because for almost every $v \in \Omega$ the operator $\left(A_{v}, D\left(A_{v}\right)\right)$ is dissipative in $X$, we have by Eq. (2.65) that for $u(v) \in D\left(A_{v}\right)$,

$$
\begin{equation*}
\left\|\left(\lambda I-A_{v}\right) u(v)\right\|_{X} \geq \lambda\|u(v)\|_{X}, \quad \lambda>0, \text { a.a. } v \in \Omega \tag{2.93}
\end{equation*}
$$

Let $f \in \mathcal{X}=L_{p}(\Omega, X)$. Because for $v \in \Omega$ we have $f(v) \in X$, by $m$ dissipativity of $A_{v}$ there is $u(v) \in D\left(A_{v}\right)$ satisfying $\left(\lambda I-A_{v}\right) u(v)=f(v)$ for almost all $v \in \Omega$ and therefore $u(v)=R\left(\lambda, A_{v}\right) f(v)$. By (2.93) we get

$$
\begin{equation*}
\|u(v)\|_{X}=\left\|R\left(\lambda, A_{v}\right) f(v)\right\|_{X} \leq \lambda^{-1}\|f(v)\|_{X} \tag{2.94}
\end{equation*}
$$

The function $u$ defined by $v \rightarrow u(v)$ is measurable by assumption and by integration we have

$$
\|u\|_{\mathcal{X}} \leq \lambda^{-1}\|f\|_{\mathcal{X}}
$$

Hence $u \in \mathcal{X}$ by Theorem ??. Consequently, again by (2.65), $\mathcal{A}$ is dissipative in $\mathcal{X}$ and $\lambda \mathcal{I}-\mathcal{A}$ is surjective onto $\mathcal{X}$. Hence $\mathcal{A}$ generates a semigroup of contractions, say $(\mathcal{G}(t))_{t \geq 0}$, in $\mathcal{X}$.

Let $\mathcal{R}(\lambda, \mathcal{A})$ be the resolvent of $\mathcal{A}$. From the preceding considerations it follows that for every $f \in \mathcal{X}$,

$$
\begin{equation*}
[\mathcal{R}(\lambda, \mathcal{A}) f](v)=R\left(\lambda, A_{v}\right) f(v) \tag{2.95}
\end{equation*}
$$

By Eq. (2.45) we have, for an arbitrary $u \in \mathcal{X}$ and $t \geq 0$,

$$
\mathcal{G}(t) u=\lim _{n \rightarrow \infty}\left(\frac{n}{t} \mathcal{R}\left(\frac{n}{t}, \mathcal{A}\right)\right)^{n} u
$$

and we can extract a subsequence $\left(\left(n_{k} \mathcal{R}\left(n_{k} / t, \mathcal{A}\right) / t\right)^{n_{k}} u\right)_{k \in \mathbb{N}}$ which converges in $\mathcal{X}$ almost everywhere in $\Omega$. On the other hand this subsequence converges in $X$ to $G_{v}(t) u(v)$, by (2.95), because $A_{v}$ is the generator of $\left(G_{v}(t)\right)_{t \geq 0}$. Therefore (2.92) holds.

Example 2.42. Let us consider the following simple transport problem. Find $f \in L_{1}\left(\mathbb{R}_{+}^{2}\right)$ satisfying

$$
\begin{aligned}
\partial_{t} f(t, x, v) & =v \partial_{x} f(t, x, v)-a(v) f(t, x, v), \quad t>0,(x, v) \in \mathbb{R}_{+}^{2} \\
f(0, x, v) & =f_{0}(x, v)
\end{aligned}
$$

where $f_{0} \in L_{1}\left(\mathbb{R}_{+}^{2}\right)$. This model can describe the motion of particles with speed $v \geq 0$, which are absorbed at the rate $a$. In this case $f$ is the density of particles at point $x$ moving with speed $v$. About $a$ we assume that it is a measurable almost everywhere positive function. By Examples 2.19 and 2.23 we see that the resolvent $R\left(\lambda, A_{v}\right)$ and semigroup $\left(G_{v}(t)\right)_{t \geq 0}$ in $L_{1}\left(\mathbb{R}_{+}\right)$are given by, respectively,

$$
\left[R\left(\lambda, A_{v}\right) f\right](s)=\frac{1}{v} e^{(\lambda+a(v)) s / v} \int_{s}^{\infty} e^{-(\lambda+a(v)) y / v} f(y) d y
$$

and

$$
\left[G_{v}(t) f\right](x)=e^{-a(v) t} f(x+v t)
$$

From, for example, the Fubini theorem $\left[R\left(\lambda, A_{v}\right) f\right](s)$ is measurable as a function of two variables and thus, by Proposition 2.41, we obtain that the semigroup for the full problem is given by

$$
[\mathcal{G}(t) f](x, v)=e^{-a(v) t} f(x+v t, v)
$$

Chapters ?? and ?? are concerned with more realistic, and certainly more involved, transport problems.

### 2.4 Nonhomogeneous Problems

Nonhomogeneous problems do not belong to the mainstream of topics that concern us in this monograph. Occasionally, however, we need some basic results. We recall them at this point.

Let us consider the problem of finding the solution to the Cauchy problem:

$$
\begin{align*}
\frac{d u}{d t} & =A u+f(t), \quad 0<t<T \\
u(0) & =u_{0} \tag{2.96}
\end{align*}
$$

where $0<T \leq \infty, A$ is the generator of a semigroup, and $f:(0, T) \rightarrow X$ is a known function.

If we are interested in classical solutions then $f$ must be continuous. However, this condition proves to be insufficient. We thus generalise the concept of mild solution introduced in (2.20). We observe that if $u$ is a classical solution of (2.96), then it must be given by

$$
\begin{equation*}
u(t)=G(t) u_{0}+\int_{0}^{t} G(t-s) f(s) d s \tag{2.97}
\end{equation*}
$$

(see, e.g., [141, Corollary 4.2.2]). The integral is well defined even if $f \in$ $L_{1}([0, T], X)$ and $u_{0} \in X$. We call $u$ defined by (2.97) the mild solution of (2.96). For an integrable $f$ such $u$ is continuous but not necessarily differentiable, and therefore it may be not a solution to (2.96).

We have the following theorem giving sufficient conditions for a mild solution to be a classical solution (see, e.g., [141, Corollary 4.2.5 and 4.2.6]).

Theorem 2.43. Let $A$ be the generator of a $C_{0}$-semigroup $(G(t))_{t \geq 0}$ and $x \in$ $D(A)$. Then (2.97) is a classical solution of (2.96) if either
(i) $f \in C^{1}([0, T], X)$, or
(ii) $f \in C([0, T], X) \cap L_{1}([0, T], D(A))$.

The assumptions of this theorem are often too restrictive for applications. On the other hand, it is not clear exactly what the mild solutions solve. A number of weak formulations of (2.96) have been proposed (see, e.g., [82, pp. 88-89] or [47]), all of them having (2.97) as their solutions. We present here a result from [79, p. 451] which is particularly suitable for our applications in Subsection ??.

Proposition 2.44. A function $u \in C\left(\mathbb{R}_{+}, X\right)$ is a mild solution to (2.96) with $f \in L_{1}\left(\mathbb{R}_{+}, X\right)$ in the sense of (2.97) if and only if $\int_{0}^{t} u(s) d s \in D(A)$ and

$$
\begin{equation*}
u(t)=u_{0}+A \int_{0}^{t} u(s) d s+\int_{0}^{t} f(s) d s, \quad t \geq 0 \tag{2.98}
\end{equation*}
$$

Proof. Suppose $u$ satisfies (2.98). Because, by assumption, $u$ is continuous, (2.98) can be written as

$$
\frac{d}{d t} \int_{0}^{t} u(s) d s=u_{0}+A \int_{0}^{t} u(s) d s+\int_{0}^{t} f(s) d s
$$

hence $\int_{0}^{t} u(s) d s$ is the solution of (2.96) with inhomogeneity $u_{0}+\int_{0}^{t} f(s) d s$ and zero initial condition. Hence, by the discussion preceding (2.97),

$$
\begin{align*}
\int_{0}^{t} u(s) d s & =\int_{0}^{t} G(t-s)\left(u_{0}+\int_{0}^{s} f(\sigma) d \sigma\right) d s \\
& =\int_{0}^{t} G(t-s) u_{0} d s+\int_{0}^{t} G(t-s)\left(\int_{0}^{s} f(\sigma) d \sigma\right) d s \\
& =\int_{0}^{t} G(s) u_{0} d s+\int_{0}^{t} G(t-s)\left(\int_{0}^{s} f(\sigma) d \sigma\right) d s \tag{2.99}
\end{align*}
$$

Now if $f \in L_{1}\left(\mathbb{R}_{+}, X\right)$, then $(s, \sigma) \rightarrow F(s, \sigma):=G(s) f(\sigma)$ is an integrable function on $[0, t] \times[0, t]$. In fact, if $f$ is a simple function, then $F$ is measurable by Example ??, because $(G(t))_{t \geq 0}$ is strongly continuous. If $f$ is only integrable, then it can be approximated by simple functions $\left(f_{n}\right)_{n \in \mathbb{N}}$ and then obviously $F_{n}(s, \sigma):=G(s) f_{n}(\sigma)$ converges to $F$ in $L_{1}([0, t] \times[0, t])$, by local uniform boundedness of $(G(t))_{t \geq 0}$. Thus, changing first the order of integration and then the variable $s$ in the inner integral according to $t-s=r-\sigma$, we get

$$
\begin{aligned}
& \int_{0}^{t} G(t-s)\left(\int_{0}^{s} f(\sigma) d \sigma\right) d s=\int_{0}^{t}\left(\int_{\sigma}^{t} G(t-s) f(\sigma) d s\right) d \sigma \\
= & \int_{0}^{t}\left(\int_{\sigma}^{t} G(r-\sigma) f(\sigma) d r\right) d \sigma=\int_{0}^{t}\left(\int_{0}^{r} G(r-\sigma) f(\sigma) d \sigma\right) d r
\end{aligned}
$$

where to get the last integral we changed the order of integration once again. Therefore (2.99) takes the form

$$
\begin{equation*}
\int_{0}^{t} u(s) d s=\int_{0}^{t} G(s) u_{0} d s+\int_{0}^{t}\left(\int_{0}^{r} G(r-\sigma) f(\sigma) d \sigma\right) d r \tag{2.100}
\end{equation*}
$$

Differentiating and taking into account that $u(t)$ and $G(t) u_{0}$ are continuous, we obtain (2.97).

To prove the converse we note that $u$, defined by (2.97), is continuous. Integrating (2.97) and performing the above calculations in the reverse order we obtain (2.99),

$$
\int_{0}^{t} u(s) d s=\int_{0}^{t} G(t-s)\left(u_{0}+\int_{0}^{s} f(\sigma) d \sigma\right) d s
$$

If $f(t)$ were continuous then we would be able to use Theorem 2.43 to claim that $v(t):=\int_{0}^{t} u(s) d s$ is a classical solution to (2.96) and then, by differentiating $v$, to obtain (2.98). For $f$ that is only integrable we have to proceed
with more care. Consider $u_{0}+F(t)=u_{0}+\int_{0}^{t} f(\sigma) d \sigma$. By (2.14) and (2.13) we obtain that $\int_{0}^{t} G(t-s) u_{0} d s \in D(A)$ and is differentiable with the derivative $G(t) u_{0}$ so that we have to show that $v_{1}(t):=\int_{0}^{t} G(t-s) F(s) d s \in D(A)$. By the definition of the domain we have to consider

$$
\begin{aligned}
\frac{G(h)-I}{h} v_{1}(t) & =\frac{1}{h}\left(\int_{0}^{t} G(t+h-s) F(s) d s-\int_{0}^{t} G(t-s) F(s) d s\right) \\
& =\frac{1}{h}\left(v_{1}(t+h)-v_{1}(t)-\int_{t}^{t+h} G(t+h-s) F(s) d s\right) .
\end{aligned}
$$

Because $u$ is continuous, $v$ is continuously differentiable and so is $v_{1}(t)=$ $v(t)-\int_{0}^{t} G(t-s) u_{0} d s$. Hence we only have to deal with the second term. Noting that, by (2.13),

$$
\lim _{h \rightarrow 0^{+}} \frac{1}{h} \int_{t}^{t+h} G(t+h-s) F(t) d s=\lim _{h \rightarrow 0^{+}} \frac{1}{h} \int_{0}^{h} G(u) F(t) d u=F(t)
$$

we obtain, using uniform continuity of $F$ to pass to the limit in the last line,

$$
\begin{aligned}
\lim _{h \rightarrow 0^{+}} & \frac{1}{h} \int_{t}^{t+h} G(t+h-s) F(s) d s \\
& =F(t)+\lim _{h \rightarrow 0^{+}} \frac{1}{h} \int_{t}^{t+h} G(t+h-s)(F(s)-F(t)) d s \\
& =F(t)+\lim _{h \rightarrow 0^{+}} \frac{1}{h} \int_{0}^{h} G(u)(F(t+h-u)-F(t)) d u=F(t)
\end{aligned}
$$

Thus, by (2.13)

$$
\begin{aligned}
A v(t) & =A \int_{0}^{t} G(t-s) u_{0} d s+\lim _{h \rightarrow 0^{+}} \frac{G(h)-I}{h} v_{1}(t) \\
& =G(t) u_{0}-u_{0}+u(t)-G(t) u_{0}-\int_{0}^{t} f(s) d s
\end{aligned}
$$

hence $v(t) \in D(A)$ for any $t>0$ and satisfies (2.98).

### 2.5 Positive Semigroups

Definition 2.45. Let $X$ be a Banach lattice. We say that the semigroup $(T(t))_{t \geq 0}$ on $X$ is positive if for any $x \in X_{+}$and $t \geq 0$,

$$
T(t) x \geq 0
$$

We say that an operator $(A, D(A))$ is resolvent positive if there is $\omega$ such that $(\omega, \infty) \subset \rho(A)$ and $R(\lambda, A) \geq 0$ for all $\lambda>\omega$.

Remark 2.46. In this section, because we address several problems related to spectral theory, we need complex Banach lattices. Let us recall, Definitions ?? and ??, that a complex Banach lattice is always a complexification $X_{C}$ of an underlying real Banach lattice $X$. In particular, $x \geq 0$ in $X_{C}$ if and only if $x \in X$ and $x \geq 0$ in $X$.

It is easy to see that a strongly continuous semigroup is positive if and only if its generator is resolvent positive. In fact, the positivity of the resolvent for $\lambda>\omega$ follows from (2.54) and closedness of the positive cone; see Proposition ??. Conversely, the latter with the exponential formula (2.45) shows that resolvent positive generators generate positive semigroups.

A number of spectral results for semigroups can be substantially improved if the semigroup in question is positive. The following theorem holds, [139, Theorem 1.4.1].

Theorem 2.47. Let $(G(t))_{t \geq 0}$ be a positive semigroup on a Banach lattice, with generator $A$. Then

$$
\begin{equation*}
R(\lambda, A) x=\int_{0}^{\infty} e^{-\lambda t} G(t) x d t \tag{2.101}
\end{equation*}
$$

for all $\lambda \in \mathbb{C}$ with $\Re \lambda>s(A)$. Furthermore,
(i) Either $s(A)=-\infty$ or $s(A) \in \sigma(A)$;
(ii) For a given $\lambda \in \rho(A)$, we have $R(\lambda, A) \geq 0$ if and only if $\lambda>s(A)$;
(iii) For all $\Re \lambda>s(A)$ and $x \in X$, we have $|R(\lambda, A) x| \leq R(\Re \lambda, A)|x|$.

Proof. From Proposition 2.29 we have $s(A) \leq \operatorname{abs}(G)=\omega_{1}(G)$, hence if $\omega_{1}(G)=-\infty$, then $s(A)=-\infty$. We may therefore assume that $\omega_{1}(G)>-\infty$. First, we prove that $\omega_{1}(G) \in \sigma(A)$. If we assume the contrary then $R(\lambda, A)$, which is analytic for $\Re \lambda>\omega_{1}(G)(=a b s(G))$, can be extended to an $\epsilon$ neighbourhood of $\omega_{1}(G)$ for some $\epsilon>0$. However, by Proposition 2.28, for $\Re \lambda>\omega_{1}(G)$ the resolvent $R(\lambda, A)$ is the Laplace transform of $t \rightarrow G(t) x$. If $x \geq 0$, then Theorem ?? implies that $\mathcal{L}(G(t) x)$ can be extended analytically for $\Re \lambda>\omega_{1}(G)-\epsilon$. By decomposing any $x \in X$ into real and imaginary parts (see Definition ??) and then each of these into positive and negative parts,
we see that $\mathcal{L}(G(t) x)$ exists for any $x \in X$ in the half-plane $\Re \lambda>\omega_{1}(G)-\epsilon$ which contradicts the result that $\omega_{1}(G)=a b s(G)$ (see Proposition 2.29). Thus, $\omega_{1}(G) \in \sigma(A)$ and therefore $\omega_{1}(G) \leq s(A)$ which yields $s(A)=\omega_{1}(G)$.

To prove (ii) we note that if $\lambda>s(A)$, then by the first part of the proof, $\lambda>a b s(G)$, and thus $R(\lambda, A)$, given by $(2.101)$, is positive as $(G(t))_{t \geq 0}$ is positive. Conversely, assume that $\lambda \in \mathbb{C}$ and $R(\lambda, A) \geq 0$. We begin by proving that $\lambda \in \mathbb{R}$. Let $x \geq 0, y=R(\lambda, A) x \geq 0$, and note that by Definition ??,

$$
\overline{A y}=\lim _{t \rightarrow 0^{+}} \frac{1}{\bar{t}} \overline{(G(t) y-y)}=\lim _{t \rightarrow 0^{+}} \frac{1}{t}(G(t) y-y)=A y
$$

and thus the identity

$$
\lambda y-A y=x=\bar{x}=\overline{\lambda y-A y}=\bar{\lambda} y-A y
$$

shows that $\lambda=\bar{\lambda}$.
Because for all $\mu \geq s(A)$ we have $R(\mu, A) \geq 0$, taking arbitrary $\mu>$ $\max \{\lambda, s(A)\}$, we obtain from the resolvent identity (2.22),

$$
\begin{equation*}
R(\lambda, A)=R(\mu, A)+(\mu-\lambda) R(\lambda, A) R(\mu, A) \geq R(\mu, A) \geq 0 \tag{2.102}
\end{equation*}
$$

by assumption. Because $\lambda \in \rho(A), \lambda \neq s(A)$ (by the first part of the proof). Suppose $\lambda<s(A)$. Then we can take $\lambda<\mu \rightarrow s(A)$ and because $s(A) \in$ $\sigma(A)$, Theorem ?? implies $\|R(\mu, A)\| \rightarrow \infty$, which contradicts (2.102). Hence $\lambda>s(A)$.

To prove (iii) we note that by (2.101) for $x \in X$ and $\Re \lambda>s(A)$,

$$
|R(\lambda, A) x| \leq \int_{0}^{\infty} e^{-\Re \lambda t} G(t)|x| d t=R(\Re \lambda, A)|x|
$$

where the integrals are understood in the improper sense.
Remark 2.48. It can be proved, [12, Proposition 5.11.2], that the statement (ii) is true for any resolvent positive operator and not only for generators of positive semigroups but the proof of this fact is much more involved.

Example 2.49. Consider the translation semigroup on $[0,1]$ discussed in Example 2.24. It is a positive semigroup satisfying $\omega_{1}(G)=-\infty$, because $G(t) f=0$ for any $f$ if $t>1$. Consequently $s(A)=-\infty$ and $\sigma(A)=\emptyset$, in agreement with Example ??.

From Theorem 2.47 we see that the spectral bound of the generator of a positive semigroup controls the growth rate of all classical solutions. However, the strict inequality $s(A)<\omega_{0}(G)$ can still occur, as was shown by Arendt; see [139, Example 1.4.4]. In this example $X=L_{p}([1, \infty)) \cap L_{q}([1, \infty)), 1 \leq p<$ $q<\infty$, and the semigroup in question is $(G(t) f)(s):=f\left(s e^{t}\right), s>1, t>0$. Its generator is $(A f)(s)=s f^{\prime}(s)$ on the maximal domain and it can be proved that $s(A)=-1 / p<-1 / q=\omega_{0}(G)$.

Interestingly enough, $s(A)=\omega_{0}(G)$ holds for positive semigroups on $L^{p}$ spaces. This was proved a few years ago by L. Weis, [168]; see also [139, Section 3.5]. The theorem for general $p$ is quite involved so we do not present it here. However, for the case $p=1$, which is most relevant for the applications described in this book, it can be proved with much less effort.

Theorem 2.50. Let $(G(t))_{t \geq 0}$ be a positive semigroup on an $A L$-space and let $A$ be its generator. Then $s(A)=\omega_{0}(G)$.

The theorem is a corollary of a general result known as the Datko theorem.
Theorem 2.51. Let $A$ be the generator of a semigroup $(G(t))_{t \geq 0}$. If, for some $p \in[1, \infty)$,

$$
\begin{equation*}
\int_{0}^{\infty}\|G(t) x\|^{p} d t<\infty \tag{2.103}
\end{equation*}
$$

for all $x \in X$, then $\omega_{0}(G)<0$.
Proof. First we show that $(G(t))_{t \geq 0}$ is bounded. There are constants $M_{1}, \omega$ for which $\|G(t)\| \leq M_{1} e^{\omega t}$. We can assume $\omega>0$, as otherwise there is nothing to prove. From (2.103) it follows that $G(t) x \rightarrow 0$ as $t \rightarrow \infty$ for any $x \in X$. Otherwise there would be $x \in X, \delta>0$, and $\left(t_{i}\right)_{i \in \mathbb{N}}$ diverging to $\infty$ such that $\left\|G\left(t_{i}\right) x\right\| \geq \delta$, where we can assume that $t_{j}-t_{j-1} \geq \omega^{-1}$. Denote $I_{i}=\left[t_{i}-\omega^{-1}, t_{i}\right]$. Then the length of each interval $I_{i}$ is $\omega^{-1}$ and they do not overlap. The increment of $\|G(t) x\|$ over each $I_{i}$ is not greater than $M_{1} e$, therefore we see that $\|G(t) x\| \geq \delta / M_{1} e$ for $t \in I_{i}$ and any $i$. Hence

$$
\int_{0}^{\infty}\|G(t) x\|^{p} d t \geq \sum_{i=0}^{\infty} \int_{I_{i}}\|G(t) x\|^{p} d t \geq\left(\frac{\delta}{M_{1} e}\right)^{p} \sum_{i=0}^{\infty} \mu\left(I_{i}\right)=\infty
$$

which contradicts (2.103). The Banach-Steinhaus theorem implies $\|G(t)\| \leq$ $M$ and the first part is proved. Next, (2.103) implies that the map $S: X \rightarrow$ $L_{p}\left(\mathbb{R}_{+}, X\right)$ given by $S x=G(t) x$ is defined on the whole $X$ and it is also closed. In fact, let $x_{n} \rightarrow x$ in $X$ and $G(\cdot) x_{n} \rightarrow f(\cdot)$ in $L_{p}\left(\mathbb{R}_{+}, X\right)$. Then there is a subsequence of $\left(x_{n}\right)_{n \in \mathbb{N}}$ such that $G(t) x_{n_{k}} \rightarrow f(t)$ for almost every $t \in \mathbb{R}_{+}$. Because $G(t) x_{n} \rightarrow G(t) x$ for any $t \geq 0$, we obtain $f(t)=G(t) x$ almost everywhere and, by continuity, for all $t$. The Closed Graph Theorem gives now

$$
\begin{equation*}
\|S x\|^{p}=\int_{0}^{\infty}\|G(t) x\|^{p} d t \leq M_{2}^{p}\|x\|^{p} \tag{2.104}
\end{equation*}
$$

Next let us take $\rho \in\left(0, M^{-1}\right)$, where $\|G(t)\| \leq M$, and define

$$
t_{x}(\rho):=\sup \{t ;\|G(s) x\| \geq \rho\|x\| \text { for } 0 \leq s \leq t\}
$$

From the first part of the proof we see that $t_{x}(\rho)$ is finite for every $x \in X$. Moreover

$$
t_{x}(\rho) \rho^{p}\|x\|^{p} \leq \int_{0}^{t_{x}(\rho)}\|G(t) x\|^{p} d t \leq \int_{0}^{\infty}\|G(t) x\|^{p} d t \leq M_{2}^{p}\|x\|^{p}
$$

hence $t_{x}(\rho) \leq t_{0}:=\left(M_{2} / \rho\right)^{p}$ and so $x \rightarrow t_{x}(\rho)$ is uniformly bounded on $X$. Taking $t>t_{0}$, we obtain

$$
\|G(t) x\| \leq\left\|G\left(t-t_{x}(\rho)\right)\right\|\left\|G\left(t_{x}(\rho)\right) x\right\| \leq M \rho\|x\|
$$

where $M \rho<1$ by the choice of $\rho$. Finally, let us fix $t_{1}>t_{0}$ and let $t=n t_{1}+s$ with $0 \leq s<t_{1}$. Then

$$
\|G(t)\| \leq\|G(s)\|\left\|G\left(n t_{1}\right)\right\| \leq M\left\|G\left(t_{1}\right)\right\|^{n} \leq M(M \rho)^{n} \leq M^{\prime} e^{-\mu t}
$$

where $\mu=-(\ln M \rho) / t_{1}$ and $M^{\prime}=\rho^{-1}$.
Proof of Theorem 2.50. Defining $\langle f, x\rangle:=\|x\|$ for $x \in X_{+}$we obtain a positive additive functional which can be extended to a bounded positive linear functional by Theorems ?? and ??. Let $\omega>a b s(G)=s(A)$ (see Theorem 2.47). Then for $x \geq 0$ and $\tau>0$, we have

$$
\left.\int_{0}^{\tau} e^{-\omega t}\|G(t) x\| d t=\left\langle f, \int_{0}^{\tau} e^{-\omega t} G(t) x d t\right\rangle \leq<f, R(\omega, A) x\right\rangle
$$

Therefore

$$
\int_{0}^{\infty} e^{-\omega t}\|G(t) x\| d t<+\infty
$$

for all $x \in X_{+}$and hence for all $x \in X$. Theorem 2.51 then implies $\|G(t)\| \leq$ $M e^{(\omega-\mu) t}$ for some $\mu>0$, hence $\omega_{0}(G)<\omega$ which yields $\omega_{0}(G) \leq s(A)$ and consequently $s(A)=\omega_{0}(G)$.

We conclude this section by briefly describing an approach of [9] which leads to several interesting results.

To fix attention, assume for the time being that $\omega<0$ (thus, in particular, $A$ is invertible and $\left.-A^{-1}=R(0, A)\right)$ and $\lambda>0$. We note the resolvent identity

$$
-A^{-1}=(\lambda-A)^{-1}+\lambda(\lambda-A)^{-1}\left(-A^{-1}\right)
$$

which can be extended by induction to

$$
\begin{equation*}
-A^{-1}=R(\lambda, A)+\lambda R(\lambda, A)^{2}+\cdots+\lambda^{n} R(\lambda, A)^{n}\left(-A^{-1}\right) \tag{2.105}
\end{equation*}
$$

Now, because all terms above are nonnegative, we obtain

$$
\sup _{n \in \mathbb{N}, \lambda>\omega}\left\{\lambda^{n}\left\|(\lambda-A)^{-n}\left(-A^{-1}\right)\right\|_{X}\right\}=M<+\infty .
$$

This is 'almost' the Hille-Yosida estimate and allows us to prove that the Cauchy problem (2.17), (2.18) has a mild Lipschitz continuous solution for $u_{0} \in D\left(A^{2}\right)$. If, in addition, $A$ is densely defined, then this mild solution is differentiable, and thus it is a strict solution (see, e.g., [9] and [12, pp. 191200]). These results are obtained by means of the integrated, or regularised, semigroups, which are beyond the scope of this monograph, so we do not enter into details of this very rich field. We mention, however, an interesting consequence of (2.105) for semigroup generation which has already found several applications, [52, 53].
Theorem 2.52. [9, 49] Let $A$ be a densely defined resolvent positive operator. If there exist $\lambda_{0}>s(A), c>0$ such that for all $x \geq 0$,

$$
\begin{equation*}
\left\|R\left(\lambda_{0}, A\right) x\right\|_{X} \geq c\|x\|_{X} \tag{2.106}
\end{equation*}
$$

then $A$ generates a positive semigroup $\left(G_{A}(t)\right)_{t \geq 0}$ on $X$ and $s(A)=\omega_{0}\left(G_{A}\right)$.
Proof. Let us take $s(A)<\omega \leq \lambda_{0}$ and set $B=A-\omega I$ so that $s(B)<0$. Because $R(0, B)=R(\omega, A) \geq R\left(\lambda_{0}, A\right)$, it follows from (2.106) and Remark ?? that

$$
\|R(0, B) x\|_{X} \geq\left\|R\left(\lambda_{0}, A\right) x\right\|_{X} \geq c\|x\|_{X}
$$

for $x \geq 0$. Using (2.105) for $B$ and taking $x=\lambda^{n} R(\lambda, B)^{n} g, g \geq 0$ we obtain, by (2.106),
$\left\|\lambda^{n} R(\lambda, B)^{n} g\right\|_{X} \leq c^{-1}\left\|R(0, B) \lambda^{n} R(\lambda, B)^{n} g\right\| \leq c^{-1}\|R(0, B) g\|_{X} \leq M\|g\|_{X}$,
for $\lambda>0$. Again using Remark ??, we can extend the above estimate onto $X$ proving the Hille-Yoshida estimate. Because $B$ is densely defined, it generates a bounded positive semigroup and thus $\left\|G_{A}(t) f\right\| \leq e^{\omega t}$. Because $\omega>s(A)$ was arbitrary, this shows that $\omega_{0}\left(G_{A}\right) \leq s(A)$ and hence we have equality.

### 2.6 Pseudoresolvents and Approximation of Semigroups

Let $A$ be a closed, densely defined operator on $X$ and $R(\lambda, A)=(\lambda I-A)^{-1}$ be its resolvent. Let us recall that if $\mu, \lambda$ are in the resolvent set $\rho(A)$, then we have the resolvent identity

$$
R(\lambda, A)-R(\mu, A)=(\mu-\lambda) R(\lambda, A) R(\mu, A)
$$

This suggests the following definition.
Definition 2.53. Let $\Delta \subset \mathbb{C}$. A family $\{J(\lambda)\}_{\lambda \in \Delta}$ of bounded linear operators on $X$ that satisfies

$$
\begin{equation*}
J(\lambda)-J(\mu)=(\mu-\lambda) J(\lambda) J(\mu), \quad \lambda, \mu \in \Delta \tag{2.107}
\end{equation*}
$$

is called a pseudoresolvent on $\Delta$.

Theorem 2.54. Let $\{J(\lambda)\}_{\lambda \in \Delta}$ be a pseudoresolvent on $\Delta \subset \mathbb{C}$.
(a) The range $\operatorname{Im} J(\lambda)$ and the kernel $\operatorname{Ker} J(\lambda)$ are independent of $\lambda \in \Delta$;
(b) $J(\lambda)$ is the resolvent of a unique densely defined closed operator $A$ if and only if $\operatorname{Ker} J(\lambda)=\{0\}$ and $\operatorname{Im} J(\lambda)$ is dense in $X$.

Proof. (a) By

$$
J(\lambda)=J(\mu)(I+(\mu-\lambda) J(\lambda))
$$

we see that $\operatorname{Im} J(\lambda) \subset \operatorname{ImJ}(\mu)$ and, interchanging $\mu$ and $\lambda$, we obtain the equality. Similarly

$$
J(\lambda)=(I+(\mu-\lambda) J(\lambda)) J(\mu)
$$

gives $\operatorname{Ker} J(\lambda) \supset \operatorname{Ker} J(\mu)$ and by symmetry we obtain the equality.
(b) It is enough to prove sufficiency. Because $\operatorname{Ker} J(\lambda)=\{0\}, J(\lambda)$ is one-to-one and we can define for some $\lambda_{0} \in \Delta$,

$$
A=\lambda_{0} I-J\left(\lambda_{0}\right)^{-1}
$$

As defined, $A$ is linear, closed, and with $D(A)=\operatorname{Im} J\left(\lambda_{0}\right)$, dense in $X$. Also, directly from the definition, $R\left(\lambda_{0}, A\right)=J\left(\lambda_{0}\right)$. For $\lambda \in \Delta$ we have

$$
\begin{aligned}
(\lambda I-A) J(\lambda) & =\left(\left(\lambda-\lambda_{0}\right) I+\left(\lambda_{0}-A\right)\right) J(\lambda) \\
& =\left(\left(\lambda-\lambda_{0}\right) I+\left(\lambda_{0}-A\right)\right) J\left(\lambda_{0}\right)\left(I+\left(\lambda_{0}-\lambda\right) J(\lambda)\right) \\
& =\left(\lambda-\lambda_{0}\right) J\left(\lambda_{0}\right)\left(I+\left(\lambda_{0}-\lambda\right) J(\lambda)\right)+I+\left(\lambda_{0}-\lambda\right) J(\lambda) \\
& =I+\left(\lambda-\lambda_{0}\right)\left(J\left(\lambda_{0}\right)-J(\lambda)-\left(\lambda-\lambda_{0}\right) J(\lambda) J\left(\lambda_{0}\right)\right)=I
\end{aligned}
$$

Similarly

$$
\begin{aligned}
J(\lambda)(\lambda I-A) & =\left(I+\left(\lambda_{0}-\lambda\right) J(\lambda)\right) J\left(\lambda_{0}\right)\left(\left(\lambda-\lambda_{0}\right) I+\left(\lambda_{0} I-A\right)\right) \\
& =\left(I+\left(\lambda_{0}-\lambda\right) J(\lambda)\right)\left(\left(\lambda-\lambda_{0}\right) J\left(\lambda_{0}\right)+I\right) \\
& =I+\left(\lambda_{0}-\lambda\right)\left(-J\left(\lambda_{0}\right)+J(\lambda)+\left(\lambda-\lambda_{0}\right) J(\lambda) J\left(\lambda_{0}\right)\right)=I
\end{aligned}
$$

so that $J(\lambda)=R(\lambda, A)$ for every $\lambda \in \Delta$. In particular, $A$ is independent of $\lambda$ and uniquely determined by $J(\lambda)$.

Corollary 2.55. Assume that $\Delta$ is an unbounded subset of $\mathbb{C}$ and $J(\lambda)$ is a pseudoresolvent on $\Delta$. Assume that there is a sequence $\lambda_{n}$ with $\left|\lambda_{n}\right| \rightarrow \infty$ such that either
(i)

$$
\begin{equation*}
\left\|\lambda_{n} J\left(\lambda_{n}\right)\right\| \leq M \tag{2.108}
\end{equation*}
$$

for some $M<+\infty$ and $\operatorname{ImJ}(\lambda)$ is dense in $X$, or (ii)

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \lambda_{n} J\left(\lambda_{n}\right) x=x \tag{2.109}
\end{equation*}
$$

for any $x \in X$.

Then $J(\lambda)$ is the resolvent of a unique densely defined closed operator $A$.
Proof. (i) It is enough to prove that $\operatorname{Ker} J(\lambda)=\{0\}$. Clearly, $\left\|J\left(\lambda_{n}\right)\right\| \rightarrow 0$ as $n \rightarrow \infty$ and, writing (2.107) as

$$
J\left(\lambda_{n}\right)-\mu J(\mu) J\left(\lambda_{n}\right)=J(\mu)-\lambda_{n} J\left(\lambda_{n}\right) J(\mu), \quad \mu \in \Delta
$$

we get

$$
\lim _{n \rightarrow \infty}\left\|\left(\lambda_{n} J\left(\lambda_{n}\right)-I\right) J(\mu)\right\|=0
$$

Therefore, if $x \in \operatorname{Im} J(\mu)$, we have

$$
\lim _{n \rightarrow \infty} \lambda_{n} J\left(\lambda_{n}\right) x=x
$$

Because $\operatorname{Im} J(\mu)$ is dense in $X$ and $\lambda_{n} J\left(\lambda_{n}\right)$ are uniformly bounded, we have this convergence on the whole $X$. Thus, if $x \in \operatorname{Ker} J(\mu)$ then, because $\operatorname{Ker} J(\lambda)$ is independent of $\lambda, \lambda_{n} J\left(\lambda_{n}\right) x=0$ for all $n$ and therefore $x=0$, which proves the assertion.
(ii) Because $\operatorname{ImJ}(\mu)$ is a linear space independent of $\mu, \lambda_{n} J\left(\lambda_{n}\right) x \in$ $\operatorname{Im} J(\mu)$ for any $x \in X$. Hence, by (2.109), $\operatorname{Im} J(\mu)$ is dense in $X$. Also, by the Banach-Steinhaus theorem, (2.109) implies (2.108); thus the assumptions of (i) are satisfied and $J(\lambda)$ is a resolvent.

The theory of pseudoresolvents is important to develop the Trotter-Kato theory for approximation of semigroups. The main result of this theory is the following theorem.

Theorem 2.56. Let $A_{n} \in \mathcal{G}(M, \omega)$. If there exists $\lambda_{0}$ with $\Re \lambda_{0}>\omega$ such that
(a) for every $x \in X$,

$$
\lim _{n \rightarrow \infty} R\left(\lambda_{0}, A_{n}\right) x=R\left(\lambda_{0}\right) x
$$

(b) the range of $R\left(\lambda_{0}\right)$ is dense in $X$,
then there exists a unique operator $A \in \mathcal{G}(M, \omega)$ such that $R\left(\lambda_{0}\right)=R\left(\lambda_{0}, A\right)$.
Moreover, if $\left(G_{n}(t)\right)_{t \geq 0}$ are semigroups generated by $A_{n}$ and $(G(t))_{t \geq 0}$ is generated by $A$, then for any $x \in X$

$$
\begin{equation*}
\lim _{n \rightarrow \infty} G_{n}(t) x=G(t) x \tag{2.110}
\end{equation*}
$$

uniformly in $t$ on bounded intervals.
Proof. We can assume that $\omega=0$. The first step is to prove that the convergence occurs for all $\lambda$ with $\Re \lambda>0$. Define $S$ to be the set of all such $\lambda$ for which $\left(R\left(\lambda, A_{n}\right) x\right)_{n \in \mathbb{N}}$ converges. Let $\mu \in S$ and expand $R\left(\lambda, A_{n}\right)$ in the Taylor series around $\mu$ :

$$
\begin{equation*}
R\left(\lambda, A_{n}\right)=\sum_{k=0}^{\infty}(\mu-\lambda)^{k} R\left(\mu, A_{n}\right)^{k+1} \tag{2.111}
\end{equation*}
$$

We know, Eq. (2.56), that

$$
\begin{equation*}
\left\|R\left(\mu, A_{n}\right)^{k}\right\| \leq M(\Re \mu)^{-k} \tag{2.112}
\end{equation*}
$$

so that the series converges in the uniform operator topology for all $\lambda$ satisfying $|\lambda-\mu| / \Re \mu<1$ and this convergence is uniform in $\lambda$ for all $\lambda$ satisfying $|\lambda-\mu| / \Re \mu \leq \theta$ for any $\theta<1$. Thus, for any $\epsilon>0$ we can find $k_{0}$ such that

$$
\left\|\sum_{k=k_{0}+1}^{\infty}(\mu-\lambda)^{k} R\left(\mu, A_{n}\right)^{k+1} x\right\| \leq \frac{\|x\| M}{\Re \mu} \sum_{k=k_{0}+1}^{\infty} \theta^{k}<\epsilon\|x\| .
$$

Next, we observe that from $R\left(\mu, A_{n}\right) x \rightarrow R(\mu) x$ it follows that $R\left(\mu, A_{n}\right)^{k} x \rightarrow$ $R(\mu)^{k} x$ for any $k$. In fact,

$$
\begin{equation*}
R\left(\mu, A_{n}\right)^{k+1} x=R\left(\mu, A_{n}\right)\left(R\left(\mu, A_{n}\right)^{k} x-R(\mu)^{k} x\right)+R\left(\mu, A_{n}\right) R(\mu)^{k} x \tag{2.113}
\end{equation*}
$$

and the statement follows by induction from the boundedness of $\left\|R\left(\mu, A_{n}\right)^{k} x\right\|$. Thus, we find $n_{0}$ such that for $n, m \geq n_{0}$ and all $k \leq k_{0}$, we have

$$
\left\|R\left(\mu, A_{n}\right)^{k+1} x-R\left(\mu, A_{m}\right)^{k+1} x\right\| \leq \epsilon\|x\|
$$

Now, for such $n, m$ we obtain

$$
\begin{aligned}
\left\|R\left(\lambda, A_{n}\right) x-R\left(\lambda, A_{m}\right) x\right\| & =\left\|\sum_{k=0}^{\infty}(\mu-\lambda)^{k}\left(R\left(\mu, A_{n}\right)^{k+1} x-R\left(\mu, A_{m}\right)^{k+1} x\right)\right\| \\
& \leq \epsilon\|x\|\left(\sum_{k=0}^{k_{0}}(\theta \Re \mu)^{k}+2\right)
\end{aligned}
$$

so that $\left(R\left(\lambda, A_{n}\right)\right)_{n \in \mathbb{N}}$ strongly converges for all $\lambda$ satisfying $|\lambda-\mu|<\theta \Re \mu$ provided $\left(R\left(\mu, A_{n}\right)\right)_{n \in \mathbb{N}}$ converges. Thus, for any fixed $\mu$ with $0<\Re \mu<\Re \lambda_{0}$, any point on the closed half-plane $\{\lambda \in \mathbb{C} ; \Re \lambda \geq \Re \mu\}$ can be reached from $\lambda_{0}$ by a finite chain of disks of radius $\theta \Re \mu$ so this half-plane is in $S$. Because $\mu$ can be fixed arbitrarily with $\Re \mu>0$, we see that $S=\{\lambda \in \mathbb{C} ; \Re \lambda>0\}$.

For every $\lambda$ with $\Re \lambda>0$ we define a linear operator $R(\lambda)$ by

$$
R(\lambda) x=\lim _{n \rightarrow \infty} R\left(\lambda, A_{n}\right) x
$$

Passing to the limit in the resolvent identity for $A_{n}$ we obtain

$$
R(\lambda)-R(\mu)=(\mu-\lambda) R(\lambda) R(\mu), \quad \Re \lambda, \Re \mu>0
$$

and therefore $R(\lambda)$ is a pseudoresolvent by (2.107). By Theorem 2.54(a) the ranges of a pseudoresolvent are independent of $\lambda$, and thus we have the density of the range of $R(\lambda)$ by (b). Also, passing to the limit in (2.112), we obtain

$$
\left\|R(\lambda)^{k}\right\| \leq M \Re \lambda^{-k}
$$

so that, in particular, for $k=1$ and real $\lambda>0$ we obtain the assumption (2.108) of Corollary 2.55 and therefore $R(\lambda)$ is the resolvent of a densely
defined closed operator $A$ with $R(\lambda)=R(\lambda, A)$ that, by the above, is the generator of a semigroup of type ( $\mathrm{M}, 0$ ).

As in the proof of Proposition ??, we note that if $t \rightarrow f(t)$ is an $X$ differentiable function taking values in the domain of the generator $T$ of a semigroup $(S(t))_{t \geq 0}$, then the function $t \rightarrow S(t) f(t)$ is differentiable with

$$
\frac{d}{d t} S(t) f(t)=S(t) f^{\prime}(t)+S(t) T f(t)
$$

Using this result we have, for any fixed $t$ and $0<s<t$,

$$
\begin{aligned}
& \frac{d}{d s} G_{n}(t-s) R\left(\lambda, A_{n}\right) G(s) R(\lambda, A) x \\
& =G_{n}(t-s) R\left(\lambda, A_{n}\right) G(s) A R(\lambda, A) x-G_{n}(t-s) A_{n} R\left(\lambda, A_{n}\right) G(s) R(\lambda, A) x \\
& =\lambda G_{n}(t-s) R\left(\lambda, A_{n}\right) G(s) R(\lambda, A) x-G_{n}(t-s) R\left(\lambda, A_{n}\right) G(s) x \\
& \quad-\lambda G_{n}(t-s) R\left(\lambda, A_{n}\right) G(s) R(\lambda, A) x+G_{n}(t-s) G(s) R(\lambda, A) x \\
& =G_{n}(t-s)\left(R(\lambda, A)-R\left(\lambda, A_{n}\right)\right) G(s) x
\end{aligned}
$$

so that, integrating the left-hand side from 0 to $t$, we get

$$
R\left(\lambda, A_{n}\right) G(t) R(\lambda, A) x-G_{n}(t) R\left(\lambda, A_{n}\right) R(\lambda, A) x=R\left(\lambda, A_{n}\right)\left(G(t)-G_{n}(t)\right) R(\lambda, A) x
$$

and finally

$$
\begin{equation*}
R\left(\lambda, A_{n}\right)\left(G(t)-G_{n}(t)\right) R(\lambda, A) x=\int_{0}^{t} G_{n}(t-s)\left(R(\lambda, A)-R\left(\lambda, A_{n}\right)\right) G(s) x d s \tag{2.114}
\end{equation*}
$$

Next consider

$$
\begin{aligned}
& \left(G_{n}(t)-G(t)\right) R(\lambda, A) x \\
& \quad=G_{n}(t) R(\lambda, A) x-G(t) R(\lambda, A) x+G_{n}(t) R\left(\lambda, A_{n}\right) x-G_{n}(t) R\left(\lambda, A_{n}\right) x \\
& \quad+R\left(\lambda, A_{n}\right) G(t) x-R\left(\lambda, A_{n}\right) G(t) x \\
& \quad=G_{n}(t)\left(R(\lambda, A) x-R\left(\lambda, A_{n}\right) x\right)+\left(R\left(\lambda, A_{n}\right)-R(\lambda, A)\right) G(t) x \\
& \quad+R\left(\lambda, A_{n}\right)\left(G_{n}(t)-G(t)\right) x=I_{1, n}(t)+I_{2, n}(t)+I_{3, n}(t)
\end{aligned}
$$

Let us fix $t_{1}<+\infty$ and let $t \in\left[0, t_{1}\right]$. Because $\left\|G_{n}(t)\right\| \leq M$, we get $\lim _{n \rightarrow \infty} I_{1, n}(t)=0$ uniformly in $t$ on $\left[0, t_{1}\right]$. Moreover, as the set $\{G(t) x ; 0 \leq$ $\left.t \leq t_{1}\right\}$ is compact in $X$, we see that $\lim _{n \rightarrow \infty} I_{2, n}(t)=0$ uniformly in $t \in\left[0, t_{1}\right]$, as in the proof of Corollary 1.27.

To estimate $I_{3, n}(t)$ we write $x=R(\lambda, A) y$ and use (2.114) to obtain

$$
\begin{aligned}
\left\|I_{n, 3}(t)\right\| & =\left\|\int_{0}^{t} G_{n}(t-s)\left(R\left(\lambda, A_{n}\right)-R(\lambda, A)\right) G(s) y d s\right\| \\
& \leq \int_{0}^{t_{1}}\left\|G_{n}(t-s)\right\|\left\|\left(R\left(\lambda, A_{n}\right)-R(\lambda, A)\right) G(s) y\right\| d s \\
& \leq M \int_{0}^{t_{1}}\left\|\left(R\left(\lambda, A_{n}\right)-R(\lambda, A)\right) G(s) y\right\| d s .
\end{aligned}
$$

The integrand converges to zero for each $s$ and can be estimated

$$
\begin{aligned}
\left\|\left(R\left(\lambda, A_{n}\right)-R(\lambda, A)\right) G(s) y\right\| & \leq\left(\left\|R\left(\lambda, A_{n}\right)\right\|+\|R(\lambda, A)\|\right)\|G(s)\|\|y\| \\
& \leq 2 M^{2} \Re \lambda^{-1}\|y\|
\end{aligned}
$$

from the Hille-Yosida theorem. Thus, by the Lebesgue dominated convergence theorem, $I_{n, 3}(t)$ also tends to zero uniformly in $t \in\left[0, t_{1}\right]$. Hence, we have

$$
\lim _{n \rightarrow \infty}\left\|R\left(\lambda, A_{n}\right)\left(G_{n}(t)-G(t)\right) R(\lambda, A) y\right\|=0
$$

uniformly in $t \in\left[0, t_{1}\right]$. Thus, for any $x=R(\lambda, A)^{2} y \in D\left(A^{2}\right)$,

$$
\begin{equation*}
\lim _{n \rightarrow \infty}\left\|\left(G_{n}(t)-G(t)\right) x\right\|=0 \tag{2.115}
\end{equation*}
$$

uniformly in $t$. Because $\left\|G_{n}(t)-G(t)\right\|$ is uniformly bounded and $D\left(A^{2}\right)$ is dense in $X$, by the Banach-Steinhaus theorem (2.115) can be extended to $X$.

Corollary 2.57. If the limit

$$
\begin{equation*}
\lim _{\lambda \rightarrow \infty} \lambda R\left(\lambda, A_{n}\right) x=x \tag{2.116}
\end{equation*}
$$

is uniform in $n$, then $R(\lambda)$ is the resolvent of a densely defined closed operator.
Proof. Writing the assumption explicitly as: for any $\epsilon$ there is $\lambda_{0}$ such that for every $\lambda>\lambda_{0}$ and any $n\left\|\lambda R\left(\lambda, A_{n}\right) x-x\right\| \leq \epsilon$, we can pass to the limit inside, so that $\|\lambda R(\lambda) x-x\| \leq \epsilon$ and condition (ii) of Corollary 2.55 is satisfied.

Theorem 2.58. Assume that $A_{n} \in G(M, \omega)$ satisfy
(i) $A_{n} x \rightarrow A x$ as $n \rightarrow \infty$ on a dense subset $D$ of $X$,
(ii) $\overline{\operatorname{Im}\left(\lambda_{0} I-A\right) D}=X$ for some $\lambda_{0}>\omega$,
then $\bar{A} \in G(M, \omega)$ and the assertions of Theorem 2.56 hold.

Proof. We begin by proving the convergence of resolvents of $A_{n}$. Let $y \in D$, $x=\left(\lambda_{0} I-A\right) y$, and $x_{n}=\left(\lambda_{0} I-A_{n}\right) y$. Because $A_{n} y \rightarrow A y$, we see that $x_{n} \rightarrow x$ as $n \rightarrow \infty$. Also

$$
\begin{align*}
\lim _{n \rightarrow \infty} R\left(\lambda_{0}, A_{n}\right) x & =\lim _{n \rightarrow \infty}\left(R\left(\lambda_{0}, A_{n}\right)\left(x-x_{n}\right)+R\left(\lambda_{0}, A_{n}\right) x_{n}\right) \\
& =\lim _{n \rightarrow \infty} R\left(\lambda_{0}, A_{n}\right)\left(x-x_{n}\right)+y=y \tag{2.117}
\end{align*}
$$

on account of the norm boundedness of $R\left(\lambda_{0}, A_{n}\right)$. Thus, $R\left(\lambda_{0}, A_{n}\right)$ converges on $\left(\lambda_{0} I-A\right) D$. But this set is dense in $X$ and again using boundedness of $\left\|R\left(\lambda_{0}, A_{n}\right)\right\|$ we obtain convergence of $R\left(\lambda_{0}, A_{n}\right)$ on the whole space. Define

$$
\begin{equation*}
\lim _{n \rightarrow \infty} R\left(\lambda_{0}, A_{n}\right) x=R\left(\lambda_{0}\right) x \tag{2.118}
\end{equation*}
$$

From (2.117) we see that $D$ is contained in the range of $R\left(\lambda_{0}\right)$ hence the latter is dense. By Theorem 2.56 we have the existence of an operator $A^{\prime} \in G(M, \omega)$ such that $R\left(\lambda_{0}, A^{\prime}\right)=R\left(\lambda_{0}\right)$.

To prove that $A^{\prime}=\bar{A}$, we first show that $A^{\prime} \supset A$. For $x \in D$ we have

$$
\lim _{n \rightarrow \infty} R\left(\lambda_{0}, A_{n}\right)\left(\lambda_{0} I-A\right) x=R\left(\lambda_{0}, A^{\prime}\right)\left(\lambda_{0} I-A\right) x
$$

and on the other hand

$$
\begin{aligned}
R\left(\lambda_{0}, A_{n}\right)\left(\lambda_{0} I-A\right) x & =R\left(\lambda_{0}, A_{n}\right)\left(\lambda_{0} I-A_{n}\right) x+R\left(\lambda_{0}, A_{n}\right)\left(A_{n}-A\right) x \\
& =x+R\left(\lambda_{0}, A_{n}\right)\left(A_{n}-A\right) x \rightarrow x
\end{aligned}
$$

as $n \rightarrow \infty$ due to the norm boundedness of $R\left(\lambda_{0}, A_{n}\right)$. Therefore

$$
R\left(\lambda_{0}, A^{\prime}\right)\left(\lambda_{0} I-A\right) x=x
$$

for $x \in D$ so that $A^{\prime} x=A x$ on $D$ and therefore $A^{\prime} \supset A$. Let $y^{\prime}=A^{\prime} x^{\prime}$ so that $\lambda_{0} x^{\prime}-A^{\prime} x^{\prime}=\lambda_{0} x^{\prime}-y^{\prime}$. Because $A^{\prime}$ is an extension of $A,\left(\lambda_{0} I-A^{\prime}\right) D$ is dense in $X$ and there is a sequence $\left(x_{n}\right)_{n \in \mathbb{N}} \subset D$ such that

$$
\lim _{n \rightarrow \infty} y_{n}=\lim _{n \rightarrow \infty}\left(\lambda_{0} I-A^{\prime}\right) x_{n}=\lim _{n \rightarrow \infty}\left(\lambda_{0} I-A\right) x_{n}=\lambda_{0} x^{\prime}-y^{\prime}
$$

Thus

$$
\lim _{n \rightarrow \infty} x_{n}=\lim _{n \rightarrow \infty} R\left(\lambda_{0}, A^{\prime}\right) y_{n}=R\left(\lambda_{0}, A^{\prime}\right)\left(\lambda_{0} x^{\prime}-A^{\prime} x^{\prime}\right)=x^{\prime}
$$

and

$$
\lim _{n \rightarrow \infty} A x_{n}=\lim _{n \rightarrow \infty}\left(\lambda_{0} x_{n}-y_{n}\right)=y^{\prime}
$$

Therefore $y^{\prime}=\bar{A} x^{\prime}$ and $A^{\prime} \subset \bar{A}$. This proves $A^{\prime}=\bar{A}$.

### 2.7 Uniqueness and Nonuniqueness

Let us return to the general Cauchy problem (2.1), (2.2). If, for a given $u_{0}$, it has two solutions, then their difference is again a solution of (2.1) but corresponding to the null initial condition - it is called a nul-solution; see [100, Section 23.7]. We say that a solution is of normal type $\omega$ if

$$
\begin{equation*}
\limsup _{t \rightarrow \infty} t^{-1} \log \|u(t)\|=\omega<+\infty \tag{2.119}
\end{equation*}
$$

A solution $u(t)$ is said to be of normal type if it is of normal type $\omega$ for some $\omega<+\infty$.
Remark 2.59. It is easy to see that if $u(t)$ is of normal type $\omega$, then for any $\omega^{\prime}>\omega$ there is $M_{\omega^{\prime}}$ such that

$$
\|u(t)\| \leq M_{\omega^{\prime}} e^{\omega^{\prime} t}
$$

Indeed, otherwise there would be $\bar{\omega}>\omega$ such that for any $n$ there would be $t_{n}$ with

$$
\left\|u\left(t_{n}\right)\right\| \geq n e^{\bar{\omega} t_{n}}
$$

We can assume that $\left(t_{n}\right)_{n \in \mathbb{N}}$ is unbounded as otherwise there would be a subsequence converging to a finite value $t$ at which the solution would blowup, contrary to the assumption that the solution is continuous for all $t$. Thus

$$
\limsup _{t \rightarrow \infty} t^{-1} \log \|u(t)\| \geq \frac{\log n}{t_{n}}+\bar{\omega}>\omega
$$

Conversely, if the solution is exponentially bounded, then it is of normal type.
Theorem 2.60. [100, Theorem 23.7.1] If $\mathcal{A}$ is a closed operator whose point spectrum is not dense in any right half-plane, then for each $u_{0} \in X$ the Cauchy problem of Definition 2.1 has at most one solution of normal type.

Proof. If there are two solutions of possibly different, normal type, then their difference, say $u$, is a nul-solution of some normal type, say $\omega$. Let

$$
\mathcal{L}(\lambda) u=\int_{0}^{\infty} e^{-\lambda t} u(t) d t
$$

where the integral exists as the Bochner integral for $\Re \lambda>\omega$ where it defines a holomorphic function. For such $\lambda$ and $0<\alpha<\beta<+\infty$ we have

$$
\int_{\alpha}^{\beta} e^{-\lambda t} u^{\prime}(t) d t=\int_{\alpha}^{\beta} e^{-\lambda t} \mathcal{A} u(t) d t=\mathcal{A} \int_{\alpha}^{\beta} e^{-\lambda t} u(t) d t
$$

where we used the closedness of $\mathcal{A}$. Integrating the first term by parts we have

$$
\int_{\alpha}^{\beta} e^{-\lambda t} u^{\prime}(t) d t=e^{-\beta \lambda} u(\beta)-e^{-\alpha \lambda} u(\alpha)+\lambda \int_{\alpha}^{\beta} e^{-\lambda t} u(t) d t
$$

and the right-hand side converges to $\lambda L(\lambda, u)$ as $\alpha \rightarrow 0^{+}$and $\beta \rightarrow \infty$ because $u(0)=0$. Thus $\mathcal{A} \int_{\alpha}^{\beta} e^{-\lambda t} u(t) d t$ also converges and because the integral converges to $\mathcal{L}(\lambda) u$, from closedness of $\mathcal{A}$ we obtain

$$
\mathcal{A L}(\lambda) u=\lambda \mathcal{L}(\lambda) u
$$

Now, $\mathcal{L}(\lambda) u$ is not identically zero as the Laplace transform of a supposedly nonzero function and, being analytic, can be equal to zero on at most discrete set of points. Thus, $\mathcal{L}(\lambda) u$ is an eigenvector of $\mathcal{A}$ for all $\lambda$ with $\Re \lambda>\omega$ except possibly for a discrete set of $\lambda$. Thus the point spectrum is dense, contrary to the assumption.

Theorem 2.61. [100, Theorem 23.7.2] Let $\mathcal{A}$ be a closed operator. The Cauchy problem (2.1), (2.2) has a nul-solution of normal type $\leq \omega$ if and only if the eigenvalue problem

$$
\begin{equation*}
\mathcal{A} y(\lambda)=\lambda y(\lambda) \tag{2.120}
\end{equation*}
$$

has a solution $y(\lambda) \neq 0$ that is a bounded and holomorphic function of $\lambda$ in each half-plane $\Re \lambda \geq \omega+\epsilon, \epsilon>0$.

Proof. The necessity follows from the previous theorem. To prove sufficiency, assume that $y_{0}(\lambda)$ is bounded and holomorphic for $\Re \lambda \geq \omega+\epsilon$ for some $\epsilon>0$. Because the solution to (2.120) can be multiplied by an arbitrary numerical function and still be a solution, we consider $y(\lambda)=(\lambda+1-\omega)^{-3} y_{0}(\lambda)$ and take the inverse Laplace transform (??),

$$
\begin{equation*}
u(t)=\frac{1}{2 \pi i} \int_{\gamma-i \infty}^{\gamma+i \infty} e^{\lambda t} y(\lambda) d \lambda, \quad \gamma>\omega \tag{2.121}
\end{equation*}
$$

Thanks to the regularising factor, the integrand is bounded by an integrable function locally uniformly with respect to $t \in(-\infty,+\infty)$. Thus it is absolutely convergent to a function continuous in $t$ on the whole real line, which satisfies the estimate

$$
\|u(t)\| \leq 2 \sup _{-\infty<r<\infty}\left\|y_{0}(\gamma+i r)\right\| e^{\gamma t}(\gamma-\omega+1)^{2}
$$

The estimate is independent of $\gamma$ due to properties of complex integration and therefore, for $t<0$, we obtain that $y(t)=0$ by moving $\gamma$ to $\infty$. From the above we also obtain that the type of $u(t)$ does not exceed $\omega$. Using closedness of $\mathcal{A}$ we obtain

$$
\mathcal{A} u(t)=\frac{1}{2 \pi i} \int_{\gamma-i \infty}^{\gamma+i \infty} e^{\lambda t} \mathcal{A} y(\lambda) d \lambda=\frac{1}{2 \pi i} \int_{\gamma-i \infty}^{\gamma+i \infty} e^{\lambda t} \lambda y(\lambda) d \lambda
$$

Due to the fact that the regularising factor behaves as $(\Im \lambda)^{-3}$, the last integral is still absolutely convergent and equals $u^{\prime}(t)$. Thus it follows that $u(t)$ is a nul-solution of type $\leq \omega$. Clearly, $u(t)$ cannot be identically zero as it has a nonzero Laplace transform $y(\lambda)$.

Similar considerations can be carried also for mild (or integral) solutions. In the present context we say that $u$ is a mild solution of (2.1), (2.2) if $u \in$ $C([0, \infty), X), \int_{0}^{t} u(s) d s \in D(\mathcal{A})$ for any $t>0$, and

$$
\begin{equation*}
u(t)=\stackrel{\circ}{u}+\mathcal{A} \int_{0}^{t} u(s) d s, \quad t>0 \tag{2.122}
\end{equation*}
$$

As in (2.19), it is clear that $U(t)=\int_{0}^{t} u(s) d s$ is a classical solution of the nonhomogeneous problem

$$
\begin{align*}
\partial_{t} U & =\mathcal{A} U+\stackrel{\circ}{u}, \quad t>0 \\
\lim _{t \rightarrow 0^{+}} U(t) & =0 \tag{2.123}
\end{align*}
$$

In particular, if $u$ is a mild nul-solution to (2.1), (2.2) of normal type $\omega$, then $U$ is a nul-solution to $(2.123)$ of the same type. We can prove the following minor modification of Theorem 2.61.

Corollary 2.62. Let $\mathcal{A}$ be a closed operator. If (2.1), (2.2) has a mild nulsolution of type $\leq \omega$, then the characteristic equation

$$
\begin{equation*}
\mathcal{A} y(\lambda)=\lambda y(\lambda) \tag{2.124}
\end{equation*}
$$

has a solution $y(\lambda) \neq 0$, which is a bounded and holomorphic function of $\lambda$ in each half-plane Re $\lambda \geq \omega+\epsilon, \epsilon>0$. Again, $y(\lambda)$ in (2.124) can be taken as

$$
\begin{equation*}
y(\lambda)=\int_{0}^{\infty} e^{-\lambda t} u(t) d t \tag{2.125}
\end{equation*}
$$

Proof. If $u$ is a mild nul-solution of type $\omega$, then $U(t)=\int_{0}^{t} u(s) d s$ is a nulsolution of the same type. Thus, by Theorem 2.61, the first part of the proposition is proved with $y(\lambda)$ of Eq. (2.124) given by

$$
y(\lambda)=\int_{0}^{\infty} e^{-\lambda t} U(t) d t
$$

Easy calculation shows that $\|y(\lambda)\|=O\left(\lambda^{-1}\right)$. Moreover,

$$
Y(\lambda):=\int_{0}^{\infty} e^{-\lambda t} u(t) d t=\lambda y(\lambda)
$$

hence $Y(\lambda)$ is a bounded holomorphic function for $\operatorname{Re} \lambda \geq \omega+\epsilon, \epsilon>0$. Because multiplication by $\lambda$ does not influence (2.124), Eq. (2.125) is proved.

Now we investigate a relation between Cauchy problems (2.1), (2.2) and (2.17), (2.18). Let $(A, D(A))$ be the generator of a $C_{0}$-semigroup $(G(t))_{t \geq 0}$ on a Banach space $X$. To simplify notation we assume that $(G(t))_{t \geq 0}$ is a semigroup of contractions, hence $\{\lambda ; \operatorname{Re} \lambda>0\} \subset \rho(A)$.

Let us further assume that there exists an extension $\mathcal{A}$ of $A$ defined on the domain $D(\mathcal{A})$. We have the following basic result.

Lemma 2.63. Under the above assumptions, for any $\lambda$ with $R e \lambda>0$,

$$
\begin{equation*}
D(\mathcal{A})=D(A) \oplus \operatorname{Ker}(\lambda I-\mathcal{A}) \tag{2.126}
\end{equation*}
$$

If we equip $D(\mathcal{A})$ with the graph norm, then $D(A)$ is a closed subspace of $D(\mathcal{A})$ and the projection of $D(\mathcal{A})$ onto $D(A)$ along $\operatorname{Ker}(\lambda I-\mathcal{A})$ is given by

$$
\begin{equation*}
x=P x^{\prime}=R(\lambda, A)(\lambda I-\mathcal{A}) x^{\prime}, \quad x^{\prime} \in D(\mathcal{A}) . \tag{2.127}
\end{equation*}
$$

Proof. Let us fix $\lambda$ with $\operatorname{Re} \lambda>0$. Because $A \subset \mathcal{A}$, then

$$
\begin{equation*}
\lambda I-A \subset \lambda I-\mathcal{A} \tag{2.128}
\end{equation*}
$$

and therefore $\operatorname{Im}(\lambda I-\mathcal{A})=X$ for $\operatorname{Re} \lambda>0$. Because $A$ is the generator of a contraction semigroup, for any $x^{\prime} \in D(\mathcal{A})$ there exists a unique $x \in D(A)$ such that

$$
(\lambda I-A) x=(\lambda I-\mathcal{A}) x^{\prime}
$$

Denote $P=R(\lambda, A)(\lambda I-\mathcal{A})$. By (2.128) it is a linear surjection onto $D(A)$, bounded as an operator from $D(\mathcal{A})$ into $D(\mathcal{A})$ equipped with the graph norm. Moreover, again by (2.128),

$$
\begin{aligned}
P^{2} & =R(\lambda, A)(\lambda I-\mathcal{A}) R(\lambda, A)(\lambda I-\mathcal{A})=R(\lambda, A)(\lambda I-A) R(\lambda, A)(\lambda I-\mathcal{A}) \\
& =R(\lambda, A)(\lambda I-\mathcal{A})=P
\end{aligned}
$$

thus it is a projection. Clearly, for $e_{\lambda} \in \operatorname{Ker}(\lambda I-\mathcal{A})$ we have $P e_{\lambda}=0$, hence this is a projection parallel to $\operatorname{Ker}(\lambda I-\mathcal{A})$. By [105, p. 155], $D(A)$ is a closed subspace of $D(\mathcal{A})$ and the decomposition (2.126) holds.

The next corollary links Theorem 2.61 with Lemma 2.63.

Corollary 2.64. If $D(\mathcal{A}) \backslash D(A) \neq \emptyset$, then $\sigma_{p}(\mathcal{A}) \supseteq\{\lambda \in \mathbb{C} ;$ Re $\lambda>0\}$. Moreover, there exists a holomorphic (in the norm of $X$ ) function $\{\lambda \in \mathbb{C} ; \operatorname{Re} \lambda>$ $0\} \ni \lambda \rightarrow e_{\lambda}$ such that for any $\lambda$ with $\operatorname{Re} \lambda>0, e_{\lambda} \in \operatorname{Ker}(\lambda I-\mathcal{A})$, which is also bounded in any closed half-plane, $\{\lambda \in \mathbb{C} ; \operatorname{Re} \lambda \geq \gamma>0\}$.
Proof. Let $u \in D(\mathcal{A}) \backslash D(A)$ and $\mathcal{A} u=f$. For any $\lambda$ with $R e \lambda>0$, denote $g_{\lambda}=\lambda u-\mathcal{A} u$ and $v=R(\lambda, A) g_{\lambda}$, then by (2.128) $e_{\lambda}^{\prime}=u-v \in \operatorname{Ker}(\lambda I-\mathcal{A})$.

A quick calculation gives

$$
\begin{aligned}
e_{\lambda}^{\prime} & =u-v=u-R(\lambda, A)(-f+\lambda u)=u-\lambda R(\lambda, A) u+R(\lambda, A) f \\
& =-A R(\lambda, A) u+R(\lambda, A) f
\end{aligned}
$$

Taking the representation $e_{\lambda}^{\prime}=u-\lambda R(\lambda, A) u+R(\lambda, A) f$ we see that because $\lambda \rightarrow R(\lambda, A)$ is holomorphic for $R e \lambda>0, \lambda \rightarrow e_{\lambda}$ is also holomorphic there. From the Hille-Yosida theorem we have the estimate $\|R(\lambda, A)\| \leq 1 / R e \lambda$ for $R e \lambda>0$. For any scalar function $C(\lambda)$, the element $e_{\lambda}=C(\lambda) e_{\lambda}^{\prime} \in$ $\operatorname{Ker}(\lambda I-\mathcal{A})$ for each $\operatorname{Re} \lambda>0$. Thus taking, for example, $C(\lambda)=\lambda^{-1}$, we obtain $e_{\lambda}$ that satisfies the required conditions.

Proposition 2.65. If for some $\lambda>0$ the null-space $\operatorname{Ker}(\lambda I-\mathcal{A})$ is closed in $X$, then $\mathcal{A}$ is closed. In particular, $\mathcal{A}$ is closed if $\operatorname{Ker}(\lambda I-\mathcal{A})$ is finitedimensional.

Proof. We know that $\mathcal{A}$ is closed if and only if $\lambda I-\mathcal{A}$ is closed, so we prove the closedness of $\lambda I-\mathcal{A}$. Let $x_{n}^{\prime} \rightarrow x^{\prime}$ and $(\lambda I-\mathcal{A}) x_{n}^{\prime} \rightarrow y$ in $X$. Operating on $x_{n}^{\prime}$ with the projector (2.127) we obtain that $x_{n}=R(\lambda, A)(\lambda I-\mathcal{A}) x_{n}^{\prime}$ converges to some $x \in D(A)$ (both in $X$ and in $D(A))$. Thus $e_{\lambda, n}=x_{n}^{\prime}-x_{n} \in \operatorname{Ker}(\lambda I-\mathcal{A})$ also converges in $X$ and, by assumption,

$$
e_{\lambda}=\lim _{n \rightarrow \infty} e_{\lambda, n} \in \operatorname{Ker}(\lambda I-\mathcal{A})
$$

Thus

$$
x^{\prime}=x+e_{\lambda}
$$

and because both $D(A)$ and $\operatorname{Ker}(\lambda I-\mathcal{A})$ are subspaces of $D(\mathcal{A})$, we have $x^{\prime} \in$ $D(\mathcal{A})$. Moreover, because $(\lambda I-\mathcal{A}) x_{n}^{\prime} \rightarrow y$ in $X$, we have $x_{n}=R(\lambda, A)(\lambda I-$ $\mathcal{A}) x_{n}^{\prime} \rightarrow R(\lambda, A) y$; thus $R(\lambda, A) y=x$ and $(\lambda I-\mathcal{A}) x=(\lambda I-\mathcal{A}) R(\lambda, A) y=$ $(\lambda I-A) R(\lambda, A) y=y$. This finally yields

$$
(\lambda I-\mathcal{A}) x^{\prime}=(\lambda I-\mathcal{A}) x+(\lambda I-\mathcal{A}) e_{\lambda}=y
$$

and $\mathcal{A}$ is closed.
Example 2.66. In this example we develop some ideas introduced in Subsection 2.3.1. Let us consider the Dirichlet problem for the heat equation

$$
\begin{align*}
\partial_{t} u & =\Delta u, \quad \text { in } \Omega, t>0 \\
\left.u\right|_{\partial \Omega} & =0  \tag{2.129}\\
\left.u\right|_{t=0} & =\stackrel{\circ}{u}
\end{align*}
$$

where $\Omega$ is a plane domain with a polygonal boundary, [93]. We consider this problem in the space $L_{2}(\Omega)$. By Theorem 2.37, the semigroup for the above problem is generated by the restriction $A_{2}$ of the distributional Laplacian to the domain

$$
D\left(A_{2}\right)=\left\{u \in \stackrel{\circ}{W}_{2}^{1}(\Omega) ; \Delta u \in L_{2}(\Omega)\right\}
$$

Because $\Omega$ is bounded, $\left(A_{2}, D\left(A_{2}\right)\right)$ is an isomorphism from $D\left(A_{2}\right)$ onto $L_{2}(\Omega)$, [93, Theorem 2.2.2.3]. Let us denote

$$
D:=\stackrel{\circ}{W}_{2}^{1}(\Omega) \cap W_{2}^{2}(\Omega)
$$

If $\Omega$ is convex, then by Theorem $2.38, D\left(A_{2}\right)=D$ and we have the maximum possible regularity. On the other hand, if the angle $\alpha$ at one corner of $\Omega$ satisfies, say, $\pi<\alpha \leq 3 \pi / 2$, then $D=\stackrel{\circ}{W}{ }_{2}^{1}(\Omega) \cap W_{2}^{2}(\Omega)$ is a proper subspace of $D\left(A_{2}\right)$ of codimension 1 ; see [93, Theorem 4.4.3.3], [19, 20]. In other words,

$$
\begin{equation*}
\operatorname{dim} L_{2}(\Omega) / A_{2}(D)=1 \tag{2.130}
\end{equation*}
$$

As in Subsection 2.3.1, we introduce the maximal operator $A_{2, \max }$ defined to be the distributional Laplacian $\Delta$ restricted to the domain

$$
D\left(A_{2, \max }\right)=L_{2,0}(\Omega, \Delta)=\left\{u \in L_{2}(\Omega) ; \Delta u \in L_{2}(\Omega), \gamma u=0\right\}
$$

where the trace $\gamma u$ is well-defined by means of Green's theorem (see, e.g., [20]). We have the following theorem, [20].
Theorem 2.67. The operator $A_{2, \max }: L_{2,0}(\Omega, \Delta) \rightarrow L_{2}(\Omega)$ is surjective and the kernel $\operatorname{Ker}\left(A_{2, \max }\right)$ in $L_{2,0}(\Omega, \Delta)$ is isomorphic to $L_{2}(\Omega) / A_{2}(D)$.

The significance of this theorem is that because the generator $A_{2}$ : $D\left(A_{2}\right) \rightarrow L_{2}(\Omega)$ is an isomorphism, $\operatorname{Ker}\left(A_{2, \max }\right)$ is not trivial by (2.130) and functions from $\operatorname{Ker}\left(A_{2, \max }\right) \subset D\left(A_{2, \max }\right)$ do not belong to $D\left(A_{2}\right)$. Therefore $D\left(A_{2, \max }\right) \neq D\left(A_{2}\right)$ and by Theorem 2.61 and Corollary 2.64 , there exist differentiable $L_{2}(\Omega)$-valued nul-solutions to (2.129).

## References

1. Y. A. Abramovich, C. D. Aliprantis, An Invitation to Operator Theory, American Mathematical Society, Providence, RI, 2002.
2. Y. A. Abramovich, C. D. Aliprantis, Problems in Operator Theory, American Mathematical Society, Providence, RI, 2002.
3. M. Abramovitz, I. A. Stegun (Eds.), Handbook of Mathematical Functions, Dover Publications, New York, 1965.
4. R. A. Adams, Sobolev Spaces, Academic Press, New York, 1975.
5. Ch. D. Aliprantis, K.C. Border, Infinite Dimensional Analysis, Springer Verlag, Berlin, 1994.
6. Ch. D. Aliprantis, O. Burkinshaw, Positive Operators, Academic Press, Orlando, FL, 1985.
7. M. Aizenman, T. Bak, Convergence to equilibrium in a system of reacting polymers, Comm. Math. Phys, 65(3), (1979), 203-230.
8. W. J. Anderson, Continuous-Time Markov Chains. An ApplicationsOriented Approach, Springer Verlag, New York, 1991.
9. W. Arendt, Resolvent positive operators, Proc. Lond. Math. Soc. (3), 54 (2), (1987), 321-349.
10. W. Arendt, Vector-valued Laplace transforms and Cauchy problems, Israel J. Math., 59 (3), (1987), 327-352.
11. W. Arendt, A. Rhandi, Perturbation of positive semigroups. Arch. Math., 56(2), (1991) 107-119.
12. W. Arendt, Ch. J. K. Batty, M. Hieber, F. Neubrander, Vector-valued Laplace Transforms and Cauchy Problems, Birkäuser Verlag, Basel, 2001.
13. O. Arino, R. Rudnicki, Phytoplankton dynamics, Comptes Rendus Biologies, 327 (11), (2004), 961-969.
14. L. Arlotti, The Cauchy problem for the linear Maxwell-Boltzmann equation, J. Differential Equations, 69 (2), (1987), 166-184.
15. L. Arlotti, A perturbation theorem for positive contraction semigroups on $L^{1}$-spaces with applications to transport equations and Kolmogorov's differential equations, Acta Appl. Math., 23(2), (1991), 129-144.
16. L. Arlotti, J. Banasiak, Strictly substochastic semigroups with application to conservative and shattering solutions to fragmentation equations with mass loss, J. Math. Anal. Appl., 293(2), (2004), 693-720.
17. L. Arlotti, J. Banasiak, F. L. Ciake-Ciake, On well-posedness of linear Boltzmann equation of semiconductor theory, submitted.
18. L. Arlotti, B. Lods, Substochastic semigroups for transport equations with conservative boundary conditions, J. Evol. Equ., accepted.
19. J. Banasiak, G.F. Roach, On mixed boundary value problems of DirichletOblique derivative type in plane domains with piecewise differentiable boundary, J. Differential Equations, 79(1), (1989), 111-131.
20. J. Banasiak, On $L_{2}$-solvability of mixed boundary value problems for elliptic equations in plane non-smooth domains, J. Differential Equations, 97(1), (1992), 99-111.
21. J. Banasiak, J. R. Mika, Asymptotic analysis of the Fokker-Planck equation of Brownian motion, Math. Models Methods Appl. Sci., 4(1), (1994), 17-33.
22. J. Banasiak, J. R. Mika, Diffusion limit for the linear Boltzmann equation of the neutron transport theory, Math. Methods Appl. Sci., 17(13), (1994), 1071-1087.
23. J. Banasiak, Asymptotic analysis of abstract linear kinetic equation, Math. Methods Appl. Sci., 19(6), 1996, 481-505.
24. J. Banasiak, Diffusion approximation of the linear Boltzmann equation with and without an external field: Analysis of initial layer, J. Math. Anal. Appl., 205(1), (1997), 216-238.
25. J. Banasiak, Mathematical properties of inelastic scattering models in linear kinetic theory, Math. Models Methods Appl. Sci., 10(2), (2000), 163-186.
26. J. Banasiak, Diffusion approximation of an inelastic scattering model in linear kinetic theory, Adv. Math. Sci. Appl., 10(1), (2000), 375-397.
27. J. Banasiak, The existence of moments of solutions to transport equations with inelastic scattering and their application in the asymptotic analysis, J. Appl. Anal., 6(2), (2000), 13-32.
28. J. Banasiak, On the hydrodynamic limit of a linear kinetic equation with dominant elastic scattering, Atti Sem. Mat. Fis. Univ. Modena, 49(1), (2001), 221-245.
29. J. Banasiak, G. Frosali, G. Spiga, Asymptotic analysis for a particle transport equation with inelastic scattering in extended kinetic theory, Math. Models Methods Appl. Sci., 8(5), (1998), 851-874.
30. J. Banasiak, L. Demeio, Diffusion approximations of a linear kinetic equation with inelastic scattering: Asymptotic analysis and numerical results, Transport Theory Statist. Phys., 28(5), (1999), 475-498.
31. J. Banasiak, G. Frosali, G. Spiga, Inelastic scattering models in transport theory and their small mean free path analysis, Math. Methods Appl. Sci., 23(2), (2000), 121-145.
32. J. Banasiak, On a diffusion-kinetic equation arising in extended kinetic theory, Math. Methods Appl. Sci., 23(14), (2000), 1237-1256.
33. J. Banasiak, On an extension of Kato-Voigt perturbation theorem for substochastic semigroups and its applications, Taiwanese J. Math., 5(1), 2001, 169-191.
34. J. Banasiak, On non-uniqueness in fragmentation models, Math. Methods Appl. Sci., 25(7), 2002, 541-556.
35. J. Banasiak, G. Frosali, G. Spiga, An interplay between elastic and inelastic scattering in models of extended kinetic theory and their hydrodynamic limits - Reference manual, Transport Theory Statist. Phys., 31(3), (2002), 187-248.
36. J. Banasiak, On well-posedness of a Boltzmann-like semiconductor model, Math. Models Methods Appl. Sci. 13(6), (2003), 875-892.
37. J. Banasiak, On multiple solutions to linear kinetic equations, Transport Theory Statist. Phys. 30(3-4), (2003), 367-384.
38. J. Banasiak, W. Lamb, On the application of substochastic semigroup theory to fragmentation models with mass loss, J. Math. Anal. Appl., 284(1), (2003), 9-30.
39. J. Banasiak, A complete description of dynamics generated by birth-anddeath problems: A semigroup approach, in: Mathematical Modelling of Population Dynamics, R. Rudnicki (Ed.), Banach Center Publications, vol. 63, 2004, 165-176.
40. J. Banasiak, Conservative and shattering solutions for some classes of fragmentation equations, Math. Models Methods Appl. Sci., 14(4), 2004, 483-501.
41. J. Banasiak, On conservativity and shattering for an equation of phytoplankton dynamics, Comptes Rendus Biologies, 337(11), (2004), 10251036.
42. J. Banasiak, M. Mokhtar-Kharroubi, Universality of substochastic semigroups: Shattering fragmentation and explosive birth-and-death processes, Discrete Contin. Dyn. Sys. B, 5(3), (2005), 524-542.
43. J. Banasiak, J. M. Kozakiewicz, N. Parumasur, Diffusion approximation of linear kinetic equations with non-equilibrium data - Computational experiments, Transport Theory Statist. Phys., accepted.
44. J. Banasiak, M. Lachowicz, A generalization of Kato's perturbation theorem in Banach lattices, submitted.
45. J. Banasiak, M. Lachowicz, M. Moszyński, Semigroups for generalized birth-and-death equations in $l^{p}$ spaces, submitted.
46. J. Banasiak, W. Lamb, On a coagulation and fragmentation equation with mass loss, submitted.
47. J. M. Ball, Strongly continuous semigroups, weak solutions, and the variation of constants formula, Proc. Amer. Math. Soc., 63(2), (1977), 370-373.
48. C. Bardos, Problèmes aux limites pour les équations aux dérivées partielles du premier ordre à coefficients réels; théorèmes d'approximation; application à l'équation de transport, Ann. scient. Éc. Norm. Sup., $4{ }^{e}$ série, t. 3, (1970), 185-233.
49. C. J. K. Batty, D. W. Robinson, Positive one-parameter semigroups on ordered Banach spaces, Acta Appl. Math., 2(3-4), (1984), 221-296.
50. R. Beals, V. Protopopescu, Abstract time-dependent transport equations, J. Math. Analy. Appl., 121(2), (1987), 370-405.
51. A. Belleni-Morante, Applied Semigroups and Evolution Equations, Clarendon Press, Oxford, 1979.
52. A. Belleni-Morante, S. Totaro, The successive reflection method in threedimensional particle transport. J. Math. Phys., 37(6) (1996), 2815-2823.
53. A. Belleni-Morante, $B$-bounded semigroups and applications, Ann. Mat. Pura Appl. (4), 170, (1996), 359-376.
54. A. Belleni-Morante, A.C. McBride, Applied Nonlinear Semigroups, John Wiley \& Sons, Chichester, 1998.
55. N. Bellomo, A. Palczewski, G. Toscani, Mathematical topics in nonlinear kinetic theory, World Scientific, Singapore, 1988.
56. N. Bellomo, M. Pulvirenti (Eds.), Modeling in Applied Sciences, Birkhäuser, Boston, 2000.
57. J. Bertoin, Self-similar fragmentations, Ann. Inst. H. Poincar Probab. Statist., 38(3), (2002), 319-340.
58. J. Bertoin, The asymptotic behavior of fragmentation processes, J. Eur. Math. Soc., 5(4), (2003), 395-416.
59. A. T. Bharucha-Reid, Elements of the Theory of Markov Processes and Their Applications, McGraw-Hill, New York, 1960.
60. F. Bouchut, F. Golse, M. Pulvirenti, Kinetic Equations and Asymptotic Theory, Series in Applied Mathematics (Paris), 4., Gauthier-Villars, Éditions Scientifiques et Médicales Elsevier, Paris, 2000.
61. H. Brézis, Analyse Fonctionelle - Théorie et Applications, Masson Editeur, Paris, 1983.
62. H. Brézis, W. Strauss, Semi-linear second order elliptic equations in $L^{1}$, J. Math. Soc. Japan, 25 (1973), 565-590.
63. M. Cai, B. F. Edwards, H. Han, Exact and asymptotic scaling solutions for fragmentation with mass loss, Phys. Rev. A, 43 (2), (1991), 656-662.
64. G. L. Caraffini, C. E. Catalano, G. Spiga, On the small mean free path asymptotics of the transport equation with inelastic scattering, Riv. Mat. Univ. Parma, (6) 1 (1998), 13-30.
65. K. M. Case, P. F. Zweifel, Linear Transport Theory, Addison-Wesley, Reading, MA, 1967.
66. C. Cercignani, The Boltzmann Equation and Its Applications, Springer Verlag, New York, 1988.
67. C. Cercignani and D. Sattinger, Scaling Limits and Models in Physical Processes, Birkhäuser, Basel, 1998.
68. S. Chapman, G. Cowling, The Mathematical Theory of Nonuniform Gases, Cambridge University Press, Cambridge, 1952.
69. F. L. Ciake-Ciake, PhD thesis, University of KwaZulu-Natal, 2005.
70. E. B. Davies, One-parameter Semigroups, Academic Press, London, 1980.
71. E. B. Davies, Heat Kernels and Spectral Theory, Cambridge University Press, Cambridge, 1989.
72. L. Demeio, G. Frosali, Diffusion approximations of the Boltzmann equation: Comparison results for linear model problems, Atti Sem. Mat. Fis. Univ. Modena, XLVI (suppl.), (1998) 653-675.
73. W. Desch, Perturbations of positive semigroups in AL-spaces, preprint.
74. P.B. Dubovskii, I. W. Stewart, Existence, uniqueness and mass conservation for the coagulation-fragmentation equation, Math. Methods Appl. Sci., 19, (1996), 571-591.
75. N. Dunford, J.T. Schwartz, Linear Operators, Part I: General Theory, John Wiley \& Sons, New York, 1988.
76. B. F. Edwards, M. Cai, H. Han, Rate equation and scaling for fragmentation with mass loss, Phys. Rev. A, 41, (1990), 5755-5757,
77. S. N. Elaydi, An Introduction to Difference Equations, Springer Verlag, New York, 1999.
78. M. A. Eliason, J. O. Hirschfelder, General collision theory treatment for the rate of bimolecular gas phase reactions, J. Chem. Phys., 30 (1959) 1426-1436.
79. K.-J. Engel, R. Nagel, One-parameter Semigroups for Linear Evolution Equations, Graduate Texts in Mathematics, Springer Verlag, New York, 2000.
80. M. Escobedo, S. Mischler, B. Perthame, Gelation in coagulation and fragmentation models, Comm. Math. Phys., 231(1), (2002), 157-188.
81. M. Escobedo, Ph. Laurençot, S. Mischler, B. Perthame, Gelation and mass conservation in coagulation-fragmentation models, J. Differential Equations, 195(1), (2003), 143-174.
82. H. O. Fattorini, The Cauchy Problem, Addison-Wesley, Reading, M, 1983.
83. J. H. Ferziger, H. G. Kaper, Mathematical Theory of Transport Processes in Gases, North Holland, Amsterdam, 1972.
84. A.F. Filippov, On the distribution of the sizes of particles which undergo splitting, Theory Probab. Appl., 6, (1961), 275-293.
85. G. Frosali, C. van der Mee, S.L. Paveri-Fontana, Conditions for runaway phenomena in the kinetic theory of particle swarms. J. Math. Phys., 30 (5), 1989, 1177-1186.
86. G. Frosali, Asymptotic analysis for a particle transport problem in a moving medium, IMA J. Appl. Math., 60 (2), (1998), 167-185.
87. G. Frosali, C. van der Mee, F. Mugelli, A characterization theorem for the evolution semigroup generated by the sum of two unbounded operators, Math. Methods Appl. Sci., 27(6), (2004), 669-685.
88. C. R. Garibotti, G. Spiga, Boltzmann equation for inelastic scattering, J. Phys. A: Math. Gen., 27(8) (1994), 2709-2717.
89. J. A. Goldstein, Semigroups of Linear Operators and Applications, Clarendon Press, New York, 1985.
90. H. Grad, Asymptotic theory of the Boltzmann equation, Physics of Fluids, 6(2) 147-181 (1963)
91. H. Grad, Rarefied Gas Dynamics, Academic Press, New York, 1962
92. W. Greenberg, V. Protopopescu, C.V.M. van der Mee, Boundary Value Problems in Abstract Kinetic Theory, Birkhäuser Verlag, Basel, 1987.
93. P. Grisvard, Elliptic Boundary Value Problems in Nonsmooth Domains, Pitman, London, 1985.
94. M. Groppi, G. Spiga, Kinetic approach to chemical reactions and inelastic transitions in a rarefied gas, J. Math. Chem., 26 (1999), 197-219.
95. B. Haas, Loss of mass in deterministic and random fragmentations, Stochastic Process. Appl., 106(2), (2003), 245-277.
96. B. Haas, Regularity of formation of dust in self-similar fragmentations, Ann. Inst. H. Poincar Probab. Statist., 40(4), (2004), 411-438.
97. P. Hartman, Ordinary Differential Equations. Wiley, New York, 1964.
98. D. Henry, Geometric Theory of Semilinear Equations, LNM 840, Springer Verlag, New York, 1981.
99. D. Hilbert, Begründung der kinetischen Gastheorie, Math. Ann., 72, 562577 (1912).
100. E. Hille, R. S. Phillips, Functional Analysis and Semi-groups, Colloquium Publications, v. 31, American Mathematical Society, Providence, RI, 1957.
101. J. Huang, B. E. Edwards, A. D. Levine, General solutions and scaling violation for fragmentation with mass loss, J. Phys. A: Math. Gen., 24, (1991), 3967-3977.
102. J. Huang, X. Guo, B.F. Edwards, A.D. Levine, Cut-off model and exact general solutions for fragmentation with mass loss, J. Phys. A: Math. Gen., 29, (1996), 7377-7388.
103. C. Jacoboni, P. Lugli, The Monte Carlo Method for Semiconductor Device Simulation, Springer Verlag, New York, 1989.
104. I. Jeon, Stochastic fragmentation and some sufficient conditions for shattering transition, J. Korean Math. Soc., 39(4), 2002, 543-558.
105. T. Kato, Perturbation Theory for Linear Operators, 2nd ed., Springer Verlag, Berlin, 1984.
106. T. Kato, On the semi-groups generated by Kolmogoroff's differential equations, J. Math. Soc. Jap., 6(1), (1954), 1-15.
107. M. Kimmel, A. Świerniak and A. Polański, Infinite-dimensional model of evolution of drug resistance of cancer cells, J. Math. Systems Estimation Control 8(1), 1998, 1-16.
108. M. Kostoglou, A. J. Karabelas, On the breakage problem with a homogeneous erosion type kernel, J. Phys. A: Math. Gen., 34, (2001), 1725-1740.
109. W. Lamb and A. C. McBride, On a continuous coagulation and fragmentation equation with a singular fragmentation kernel, in: Evolution Equations, G. R. Goldstein, R. Nagel and S. Romanelli (Eds.), Lecture Notes in Pure and Applied Mathematics, Marcel Dekker, New York, Basel (2003), 281-298.
110. W. Lamb, Existence and uniqueness results for the continuous coagulation and fragmentation equation, Math. Methods Appl. Sci., 27(6), (2004), 703-721.
111. Ph. Laurençot, S. Mischler, The continuous coagulation-fragmentation equations with diffusion, Arch. Ration. Mech. Anal., $\mathbf{1 6 2 ( 1 ) , ~ ( 2 0 0 2 ) , ~ 4 5 - ~}$ 99.
112. A. M. Liapunov, Stability of Motion, Ph.D. thesis, Kharkov, 1892, English translation, Academic Press, New York, 1966.
113. J. C. Light, J. Ross, K. E. Shuler, Rate coefficients, reaction cross sections, and microscopic reversibility, in: Kinetic Processes in Gases and Plasmas, A. R. Hochstim, (Ed.), Academic Press, New York, 1969.
114. B. Lods, Semigroup generation properties of streaming with noncontractive boundary conditions, Mathematical and Computer Modelling, (2005), in press.
115. A. Lunardi, G. Metafune, On the domains of elliptic operators in $L^{1}$, Differential and Integral Equations, 17(1-2), (2004), 73-97.
116. W.A.J. Luxemburg, A.C. Zaanen, Riesz Spaces, North-Holland, Amsterdam, 1971.
117. A. Majorana, Space homogeneous solutions of the Boltzmann equation describing electron-phonon interaction in semiconductors, Transport Theory Statist. Phys., 20(4), (1991), 261-279.
118. A. Majorana, S.A. Marano. Space homogeneous solutions to the Cauchy problem for semiconductor Boltzmann equations, SIAM J. Math. Anal., 28(6), (1997), 1294-1308.
119. A. Majorana, C. Milazzo, Space homogeneous solutions of the linear semiconductor Boltzmann equation, J. Math. Anal. Appl., 259(2), (2001), 609-629.
120. P. A. Markowich, Ch. A. Ringhofer, Ch. Schmeiser, Semiconductor Equations, Springer Verlag, New York, 1990.
121. P. A. Markowich, Ch. Schmeiser, Relaxation time approximation for electron-phonon interaction in semiconductors, Math. Models Methods Appl. Sci. 5(4), (1995), 519-527.
122. P. A. Markowich, Ch. Schmeiser, The drift-diffusion limit for electronphonon interaction in semiconductors, Math. Models Methods Appl. Sci., 7(5), (1997), 707-729.
123. E. D. McGrady, R. M. Ziff, "Shattering" transition in fragmentation, Phys. Rev. Lett., 58(9), (1987), 892-895.
124. D. J. McLaughlin, W. Lamb, A. C. McBride, A semigroup approach to fragmentation models, SIAM J. Math. Anal., 28 (5), (1997), 1158-1172.
125. D. J. McLaughlin, W. Lamb, A. C. McBride, An existence and uniqueness theorem for a coagulation and multiple-fragmentation equation, SIAM J. Math. Anal., 28 (1997), 1173-1190.
126. C.V.M. van der Mee, Well-posedness of stationary and time-dependent Spencer-Lewis equations modelling electron slow down, J. Math. Phys. 30(1), (1989), 158-165.
127. C.V.M van der Mee, Time-dependent kinetic equations with collision terms relatively bounded with respect to the collision frequency, Transport Theory Statist. Phys., 30(1), (2001), 63-90.
128. I.V. Melnikova, A. Filinkov, Abstract Cauchy Problems: Three Approaches. Chapman \& Hall/CRC, Boca Raton, FL, 2001.
129. J. R. Mika, Singularly perturbed evolution equations in Banach space, J. Math. Anal. Apl., 58(1), (1977), 189-201.
130. J. R. Mika, New asymptotic expansion algorithm for singularly perturbed evolution equations, Math. Methods Appl. Sci., 3(2), (1981), 172-188.
131. J. R. Mika, J. Banasiak, Singularly Perturbed Evolution Equations with Applications to Kinetic Theory, Series on Advances in Mathematics for Applied Sciences, vol. 34, World Scientific, Singapore, 1995.
132. I. Miyadera Nonlinear Semigroups, American Mathematical Society, Providence, RL, 1991.
133. S. Mizohata, The Theory of Partial Differential Equations, Cambridge University Press, Cambridge, 1973.
134. M. Mokhtar-Kharroubi, Mathematical Topics in Neutron Transport Theory. New Aspects, Series on Advances in Mathematics for Applied Sciences, 46, World Scientific, River Edge, NJ, 1997.
135. F. A. Molinet, Existence, uniqueness and properties of the Boltzmann equation for a weakly ionized gas. I, J. Math. Phys., 18(5) (1977), 984996.
136. R. Nagel (Ed.), One-parameter Semigroups of Positive Operators, LNM 1184, Springer Verlag, Berlin 1986.
137. R. Nagel, Order in pure and applied functional analysis, Atti Sem. Mat. Fis. Univ. Modena, 40(2), (1992), 503-517.
138. R. Narasimhan, Analysis on Real and Complex Manifolds, North-Holland, Amsterdam, 1968.
139. J. van Neerven, The Asymptotic Behaviour of Semigroups of Linear Operators, Operator Theory: Advances and Applications 88, Birkhäuser Verlag, Basel, 1996.
140. J. R. Norris, Markov Chains, Cambridge University Press, Cambridge, 1998.
141. A. Pazy Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer Verlag, New York, 1983.
142. B. Perthame, L. Ryzhik, Exponential decay for the fragmentation or celldivision equation, J. Differential Equations, 210(1), (2005), 155-177.
143. F. Räbiger, R. Schnaubelt, A. Rhandi, J. Voigt, Non-autonomous Miyadera perturbations, Differential Integral Equations, 13(1-3), (2000), 341-368.
144. G. E. H. Reuter, W. Lederman, On differential equations for the transition probabilities of Markov processes with denumerably many states, Proc. Cambridge. Phil. Soc., 49, (1953), 247-242.
145. G. E. H. Reuter, Denumerable Markov processes and the associated contraction semigroup, Acta Math. 97, (1957), 1-46.
146. L. M. Ricciardi, Stochastic population theory: birth and death processes, in: Mathematical Ecology, T.G. Hallamand and S.A. Levin (Eds.), Springer Verlag, Berlin, 1986, 155-190.
147. A. Rossani, G. Spiga, Kinetic theory with inelastic interactions, Transport Theory Statist. Phys. 27, 273-287 (1998).
148. A. Rossani, G. Spiga, A note on the kinetic theory of chemically reacting gases, Physica A 272, 563-573 (1999).
149. H. L. Royden, Real Analysis, 2nd ed., Macmillan, New York, 1968.
150. W. Rudin, Real and Complex Analysis, 3rd ed., McGraw-Hill, New York, 1987.
151. H. H. Schaefer, Banach Lattices and Positive Operators, Springer Verlag, Berlin and New York, 1974.
152. H. H. Schaefer, On positive contractions in $L^{p}$-spaces, Trans. Amer. Math. Soc., 257(1), (1980), 261-268.
153. G. R. Sell, Y. You, Dynamics of Evolutionary Equations, Springer Verlag, New York, 2002.
154. E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, Princeton, NJ, 1970.
155. A. Świerniak, A. Polański, M. Kimmel, A. Bobrowski, J. Śmieja, Qualitative analysis of controlled drug resistance model - Inverse Laplace and semigroup approach, Control Cybernetics 28(1), (1999), 61-73.
156. H. Tanabe, Equations of Evolutions, Pitman, London, 1979.
157. H.R. Thieme, Positive perturbation of operator semigroups: Growth bounds, essential compactness, and asynchronous exponential growth, Discrete Contin. Dynam. Systems, 4(4), (1998), 735-764.
158. H.R. Thieme, Balanced exponential growth for perturbed operator semigroups, Adv. Math. Sci. Appl., 10(2), (2000), 775-819.
159. C. Truesdell, R. G. Muncaster, Fundamentals of Maxwell's Kinetic Theory of a Simple Monoatomic Gas, Academic Press, New York, 1980.
160. M. Tsuji, On Lindelöf's theorem in the theory of differential equations, Japanese J. Math., XVI, (1940), 149-161.
161. J. Voigt, On substochastic $C_{0}$-semigroups and their generators, Transport Theory Statist. Phys., 16 (4-6), (1987), 453-466.
162. J. Voigt, On the perturbation theory for strongly continuous semigroups, Math. Ann., 229, (1977), 163-171.
163. J. Voigt, Functional analytic treatment of the initial boundary value problem for collisionless gas, Habilitationschrift, Fachbereich Mathematik, Ludwig-Maximilians-Universität, München, 1980.
164. J. Voigt, On resolvent positive operators and positive $C_{0}$-semigroups on AL-spaces, Semigroup Forum, 38(2), (1989), 263-266.
165. M. Volpato, Sul problema di Cauchy per una equazione lineare alle derivative parziali del primo ordine, Rend. Sem. Mat. Univ. Padova, 28, 1958, 153-187.
166. I.I. Vrabie, Compactness Methods for Nonlinear Evolutions, Longman Scientific \& Technical, Harlow, 1987.
167. C. S. Wang Chang, G. E. Uhlenbeck, J. De Boer, The heat conductivity and viscosity of polyatomic gases, in: Studies in Statistical Mechanics, Vol. 2, J. De Boer (Ed.), North Holland, Amsterdam, 1969.
168. L. Weis, A short proof for the stability theorem for positive semigroups on $L_{p}(\mu)$, Proc. Amer. Math. Soc. 126(11) (1998), 3253-3256.
169. A. Zabczyk, A note on $C_{0}$-semigroups, Bull. Acad. Polon. Sci., 23(8), (1975), 895-898.
170. R. M. Ziff, E. D. McGrady, The kinetics of cluster fragmentation and depolymerization, J. Phys. A: Math. Gen. 18, (1985), 3027-3037.
171. R. M. Ziff, E. D. McGrady, Kinetics of polymer degradation, Macromolecules, 19 (1986), 2513-2519.
172. K. Yosida, Functional Analysis, 5th Ed., Springer Verlag, Berlin, 1978.
